Evol. Inst. Econ. Rev. 11(2): 105–120 (2014)

ARTICLE

Proactive News Article Summarization Service Using Personal Intention Models

In Seok Oh*, Ji Eun Lee** and Kyung Joong Kim***

Deptment of Computer Science and Engineering, Sejong University; Seoul, Republic of Korea.
* E-mail: ohinsuk@naver.com
** E-mail: apple173@naver.com
*** E-mail: kimkj@sejong.ac.kr

Abstract

Nowadays, we live with a huge amount of data. By IDC report, the amount of data generated in 2011 is about 1.8 ZB (trillion GB). Ironically, there are small amount of useful information when you are looking at a number of papers, internet articles, movies, pictures, and social network posts. As a result, it is required that users extract useful information manually. However, this is tedious and hard task. To solve this problem, a lot of sophisticated techniques have been proposed to provide summarization service. Although each user has different desire on the summarization, the service uses much computational resource to produce standard summarization for all users. In this study, we propose to use machine learning to predict the intention of users on the summarization. It can reduce the computational cost to summarize all the documents and make available proactive summarization service. To validate our proposal, we run experiments with eight participants for two news portals on five topic areas.

Keywords: user intention, prediction, machine learning, summarization.

1. Introduction

Summarization is one of the most important tools to fit the large amount of data into small-size personalized devices (Salton *et al*., 1997). It extracts some important sentences from the original documents and provides them as a summarization. In the process, it is necessary to rank the relevance of each sentence based on the predefined criteria. To increase the user's satisfaction, there have been some studies on the introduction of personal models for the summarization 0. In the approach, the system has a built-in module to model users' preference and exploits them to personalize the summarization algorithm for each user. It could adjust the selection of sentences based on user preference and the amount of information lost in the summarization.

JEL: C11, C30, Y10.

In the standard summarization service, they assumed that users want to get the "summarization" service instead of accessing the full contents. Because we have little knowledge on the human's cognitive process on the intention to get summarization service, it is not easy to design our service as proactive as possible. In this study, we propose to predict user's intention on the summarization and the degree of information loss. To make the prediction, our system utilizes the structural property of the documents, topic of the contents and the inclusion/exclusion of words. It is interesting to analyze which parts of the contents have big impact on the user's desire to get the summarization service. The prediction utilizes machine learning based on user's explicit feedback on some initial documents whether he/she wants to get summarization service or not.

There have been some works on relationships between human cognition and summarization services. For example, Nenkova *et al.* (2005) categorized the human's cognitive status on new articles into Hearer-Old vs. Hearer-New and Major vs. Minor. If the contents are new to the users, the summarization should provide enough details to them. If not, it is reasonable to decrease the length of the final summarization. In the second categorization, it intends to know the level (major or minor) of the characters in the story. Based on the types of characters in the news, the summarization should be produced in a different way. In the study, they used machine learning to build a model to predict the user's cognitive status from explicit feedback. In this way, they attempted to produce better summarization for human users.

Unlike previous works on the modeling human summarizers (Nenkova *et al.*, 2005; Endres-Niggemeyer and Wansorra, 2004), we attempt to infer the users' intention (or desire) to get the summarization service. Simply, it is possible to estimate that the user might want to get summarization because of the length of the article is too long. Although the hypothesis is reasonable, it's too simple to explain complex decision making of human users and there could be a lot of different factors to motivate users execute summarization service. If we ask users whether he/she wants to get automatic summarization service on some articles or not, they could answer to the questions. However, it is not easy to ask them build some rules to explain their mechanism to make decisions on the questions.

In this study, we attempt to learn user's intention on automatic summarization service on news articles. Although the proposed method can be used for other types of documents and contents, it is easy to apply to news portal because of their uniform styles of presentation. In our early work (Lee *et al.*, 2013), we conducted small-size experiments on the issue but it has been significantly expanded in this work. During navigating the news article, the user could express his/her intention on the execution of automatic summarization service. The system continuously records user's feedback and the news contents and extracts useful features from them (structural property, topic, and contents). In the learning phase, the machine learning algorithm is adopted to build user intention model automatically from

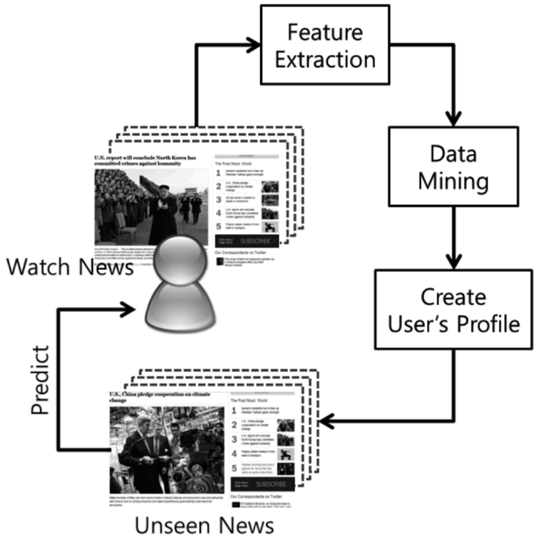


Fig. 1. The prediction on the user's intention for the automatic summarization service.

the labeled news articles. Finally, the system predicts user's intention on unseen news articles before he/she reads the article and proactively applies adequate summarization service (Fig. 1).

In the learning models for users, we collect samples from multiple users and train a single model for all users and intend to extract common features shared by users (Integrated Model). If this is successfully working, it means that there might be a single common mechanism for humans to make summarization or not. Also, we build models individually using the data for each user (Individual Model). If this is better than the integrated approach, it means that each person has different mechanism to call the automatic summarization service. Finally, we build the model separately for each news topic (Topic Model). For example, all the labeled data for politics are collected from multiple users and used for training models. It reveals the relationships between the topics and the needs on the automatic summarization.

2. Backgrounds

2.1 Cognitive modeling using machine learning

There have been several works on machine learning approaches for cognitive tasks. For example, Horvitz *et al.* (2004) collected calendar events of five participants from Microsoft Outlook messaging and appointment management system. In the study, they reviewed all the meeting events and categorized them into one of "memorable" and "non-memorable." After then, they learnt a Bayesian network to infer the probability of "memory landmark" given the evidence of events. Similarity, Kim *et al.* (2007) used Key Graph to discover chance events from mobile contents. Cho *et al.* (2007) identified the landmark events from mobile life logs collected from smart phones.

In the cognitive modeling, each participant provides with cognitive states on each data sample. After then, the machine learning system extracts useful features from the original data and builds models to correlate the features and the user's explicit feedback. Although the probabilistic models are useful to handle uncertainty in the cognitive modeling, it takes much time to predict the final outcome if the network size is big. Because of the computational cost, the Bayesian network model is not embedded into the mobile device directly. In the design of the system, it is interesting to know the inter-operability of personalized prediction models and performance degradation with the integrated model.

Summarization is human's cognitive ability to make a summary of the original contents. It can save much time to grasp important information from a massive database or records. However, it is still challenging to generate human-style summarization (abstraction) from the original documents. Instead, the traditional system attempts to select some of the important sentences from the original documents and combine them to produce the final summary. In the approach, the definition of ranking mechanism is a key to get good result. Recently, there are some works to analyze human's cognitive patterns on making summary from multiple documents. For patient records, researchers extract some important patterns on the summarization process by clinicians (Reichert *et al.*, 2010). Although they don't build practical systems, they provide important insight on human's process of summarization.

2.2 Summarization techniques

Text summarization is the one of the most important topics in natural language processing. Sometimes, we can regard the text summarization system as a text-to-text system 0. This system outputs text shorter than imputed text. This system

Table 1. Summary of automatic summarization techniques.

Author	Year	Description
R. Barzilay, M. Elhada	1997	Topic, representation (Lexical Chain)
C. Lin, E. Hovy	2000	Topic, representation (Topic Signature)
Y. Gong, X. Liu	2001	Topic, ranking method (Latent Semantic Analysis)
G. Erkan, D. R. Radev	2001	Representation (Stochastic Graph)
L. Antiqueira, O. N. Oliveira Jr. *et al.*	2009	Representation (Complex Network)
M. A. Fattah, F. Ren	2009	Selection (various Machine Learning techniques)
R. M. Aliguliyev	2009	Sentence similarity measure

consists of three parts (1) transform inputted text to intermediate representation, (2) score each sentence and (3) select importance sentences. We can analyze many works in this frame. In most of cases, each work proposes new representation, new scoring method and/or new selection methods. In the early days, many researchers find a topic in given text and score by its importance (Barzilay and Elhadad, 1997; Lin and Hovy, 2000; Gong and Liu, 2001). But, nowadays, many works use indicator instead topic (Erkan and Radev, 2004; Antiqueira *et al.*, 2009; Fattah and Ren, 2009; Aliguliyev, 2009). In this approach, they compare the importance of each sentence directly, instead of searching for the topic or interpreting the sentences. Table 1 is the summary of related works. This table summarized authors, published year and their main contribution.

2.3 Personalized news service

There are several works on the personalization of the news services. For example, Google news built profiles of user's news interests based on their past click behavior 0. In the work, they predicted user's current news interest from the activities of the user and news trends demonstrated in the activity of all users. Abel *et al.* used the activities on Twitter (popular micro-blogging site) to build a user model for personalized news recommendation system 0. Li *et al.* modeled the personalized recommendation of news articles as a contextual bandit problem and applied it to a Yahoo! Front Page Today Module dataset containing over 33 million events.

Although there have been several works on modeling human summarizer's cognitive process using machine learning, there are little works on the prediction of user's intention on the automatic summarization service. The news portals have developed algorithms to provide personalized news recommendation based on user's activities. In their works, the focus is to select the news articles in the interest of users. In this paper, we focus on the fundamental issue on the prediction of user's intention on automatic summarization service. The user intention is likely to be dependent on several factors (the amount, quality, presentation, user's context, and so on). Because of its complexity, it is not trivial to make a set of rules to capture the human's cognitive process. In this study, we build a cognitive model on the user's intention and use them to make an adaptive summarization system.

3. User Study on the Summarization Service

In this paper, we assume that the user will have different reasons and factors on the automatic summarization service. If the summarization is activated just because of the length of contents is longer than a specific threshold value, the prediction might be so simple to learn from data. To proceed on the machine learning approach, we design a user study to know user's opinion on the summarization service. In this study, 80 participants (50 males and 30 females) do the simple survey on the automatic summarization. For the user survey, we attempted to increase the diversity of participants although the number of participants was not too big. It gave us lots of hints on the design of features for the intention modeling. Its purpose is to see the reason behind user's intention on using the automatic summarization service. Their age is ranged from 20–25 (15%), 25–30 (61%) and over 30 (24%). Their job is students (63%) and employee (34%). The first question is "what is your favorite news category when you read news from online site (multiple answers allowed)?" Figure 2 shows that each user votes two or three categories in average. Also, there is no dominance of specific topics among them. They have relatively distributed interest over different topics.

Table 2 shows participants' evaluation on the possible reasons for news summarization. There might be a lot of different reasons if users explicitly express that they want the summarization for news articles. If the reason is simply the news is too long, the prediction might be relatively simple. To know the factors behind the human cognition, the user survey asks a lot of document features (structural, content and external properties) on their relevance to explicitly run document summarization service. When you make decisions on the use of automatic summarization service for news articles, is the feature important? The answer sheet has five scales (strong negative, negative, neutral and positive and strong positive). The features are sorted based on the percentage of (strong) positive.

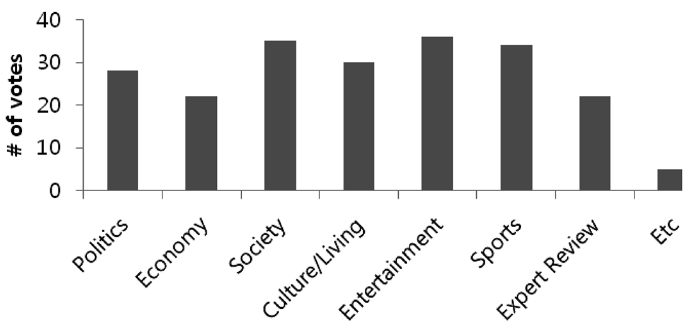


Fig. 2. The distribution of user preference on news topics.

Table 2. User evaluation on possible reasons for news summarization (percentage %).

	(Strong) Negative	Neutral	(Strong) Positive
Length of Sentences	25	12	62
Source of News	15	25	60
# of Paragraphs	20	21	58
# of Characters	23	18	57
# of Sentences	23	20	56
Familiarity on the Contents	25	26	48
Too Short or Long Title	28	23	47
Picture or Photo	52	20	27
Average	26	21	52

Participants think that the length of sentences, # of paragraphs, # of characters and # of sentences are important reasons to get the automatic summarization service. However, the positive ratio is about 60% and others are neutral or negative. That's interesting that the structural property is not obvious precursor to the summarization service. The source of news (what is the company who posts the news?) is also important consideration to get the summarization service. The length of the title is relatively weak feature to motivate user to run the summarization service. Although we assume that the existence of photos or figures could have impact on the decision making, 72% of users are neutral or negative on the usefulness of the features.

4. Applying Machine Learning to Predict User Intention on Summarization

From the user survey, it is revealed that the structural property of news is important but not dominant factor to get the summarization service. Because the design of the expert systems for the prediction of user intention is not a trivial task, machine learning approach is adopted to automate the intention modeling. At first, the system collects a set of news articles from internet news services and request users to review them. In this step, it is not necessarily read the news contents and the user needs to give explicit feedback whether he/she wants to get summarization service on the contents. While user browses the news articles, a small pop-up window will be on top of the screen. The task of user is to indicate whether the summarization is required or not.

After the labeling step, the system does a preprocessing on the news articles. Because the news is formatted in HTML and has a lot of advertisement content, it is necessary to filter out unnecessary contents from the original news. At first, the HTML tag should be removed from the source file. This process can be tailored to each news portal's style. Unfortunately, there is no general rule to be used for all the news portals. In the top, right and left sides of each news web page, there are advertisement contents. In this study, we attempt to delete the ads from the news page automatically producing pure news contents. In case that the automatic preprocessing is not successful, we manually processed some unsuccessful purification.

The next step is to extract features from the preprocessed news texts. First of all, we can count on the number of words, structure, paragraph, and sentences. Also, the length of each component could be used to guess the structure of news. The news article is divided into title and main body. It is possible to separate the title and body part when the system calculates the features. In addition to the structural property, we can count the number of photos and pictures although they're not relevant in the user survey. In addition to the structural property, the contents might be also important and the inclusion/exclusion of specific words might be important. Finally, the news topic category is extracted from the news portal.

In the learning stage, there are three ways to build user intention model for

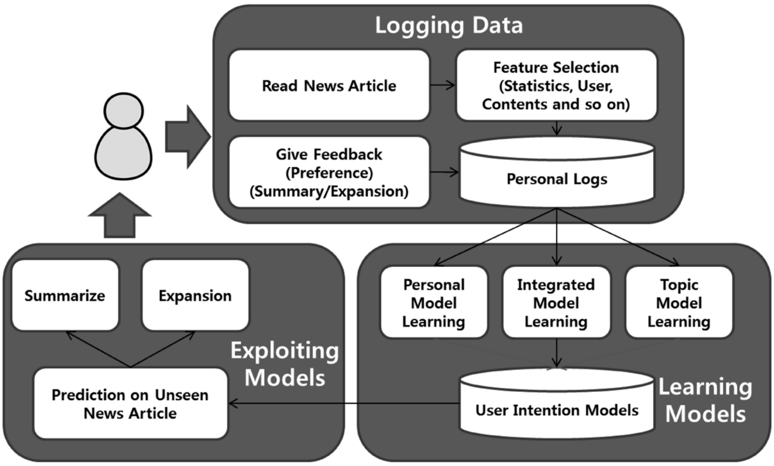


Fig. 3. Overview of learning user models to predict "summarization" intention for news articles.

summarization. 1) Integrated Model: In this method, all the labeled data from users will be integrated and used to train a single intention model 2) Individual Model: Instead of mixing all the labeled data into single training set, this method independently learns an intention model from each user's data. 3) Topic Model: It builds the intention model for each news topic. In this study, there are five news categories (Politics, Economy, Society, World and IT/Science). The goal of the integrated and topic modeling is to see the common properties on the intention prediction. The final step is to make predictions on unseen news articles (without user explicit feedback) whether he/she wants the summarization service or not. Figure 3 shows the overview of the proposed learning approach.

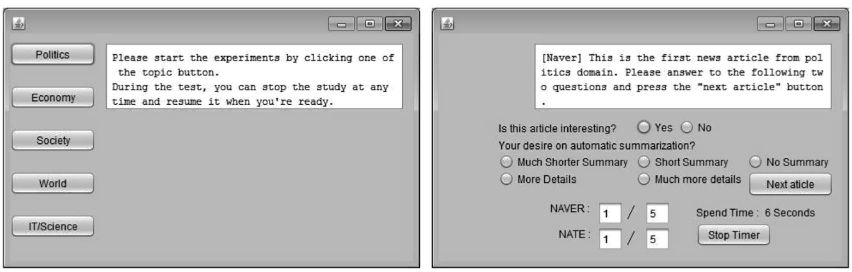
4.1 Data collection

In this study, we used two popular Korean news portal sites (Naver.com and Nate.com). From each site, 20 news articles are extracted for each topic and in total, 200 news articles are used for the user review. Figure 4 shows an example of news article extracted from the two news sites. It shows that the two news portals have different formatting styles. From the preprocessing, the system extracts the title and the main body of the news. The first step of the review is to select the news topic to review (Fig. 5). For each topic, there are 40 news articles waiting for user reviews.

The system requests each user to review the forty news articles one by one by giving explicit feedback on the needs of summarization service. It is not necessarily to read all the news contents before users give explicit feedback. Because there is no time constraint on the explicit feedback, users can use different amount of time for each news article. Eight participants are invited to review the news articles. For the user testing, we focus on the system's performance for users with diverse

(a) Naver News (b) Nate News

Fig. 4. An example of news article from the news portal site.



(a) Selection of topic (b) Explicit feedback interface

Fig. 5. User interface to get the explicit feedback from users.

interest. Although the number of participants is small, they showed quite different decision mechanism on the summarization task. In average, the review on the 200 news articles takes 65 minutes. In the feedback interface, there are five scale (much shorter summary, short summary, no summary, more details, and much more details). If users feel that the news has little information and needs expansion instead of the summarization, they can select the "details" options. To see the correlation between user's preference on the news and the summarization activation, there is one additional question on the user's preference (Is this article interesting?).

Table 3 summarizes the user feedback data. It shows that about 64% of news articles are selected as "preferable." Half of the news articles are evaluated that there is no need on the summarization. However, 30% news articles are categorized into "summarization" and 15% for "expansion." That's interesting that there are some requests on the "information expansion service." To make decision, users usually use about 20 seconds. Figure 6 shows the comparison of user intention feedback for preferable/non-preferable news contents. It shows that if the news is preferable to the user, there is more chance that the selection is "expansion." However, the difference between the two distributions is not so big.

Table 3. Summary of user feedback data.

	Preference Like (%)	Intention (%)			Time Spent (Seconds)
		Summarization	Neutral	Expansion	
User 1	59	44	38	17	8.22
User 2	74	25	51	23	24.97
User 3	72	16	74	10	12.50
User 4	63	5	69	25	15.56
User 5	75	30	42	27	45.49
User 6	66	45	51	3	11.63
User 7	57	67	33	0	13.62
User 8	53	12	71	17	24.80
Average	64	31	54	15	19.59

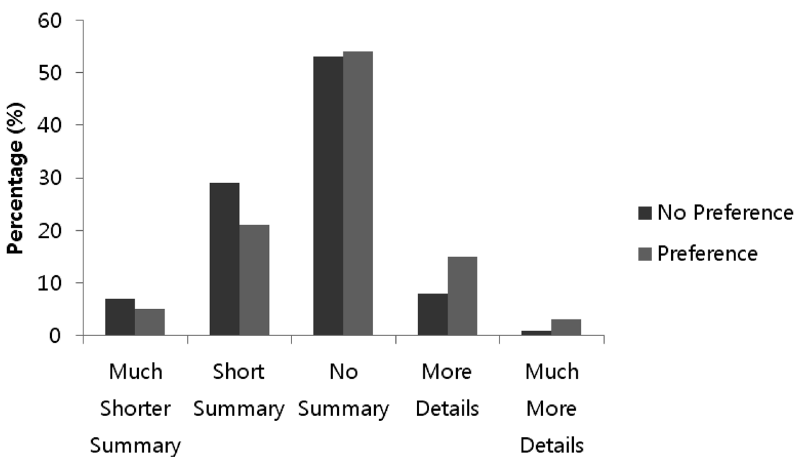


Fig. 6. The user intention distribution for preferable/non-preferable news articles.

4.2 Feature extraction

From the original news article, it simply counts the number of characters, words, sentences and paragraphs. It is possible to extract the features from the title and the main body sections independently. Although the counting can be used to estimate the structural property of the news article, it is also important to know the inclusion/exclusion of specific words in the news. From the 200 news articles, all the words are sorted based on the number of occurrence. In the selection, we only consider the noun words. In the experiments, the most popular 40 noun words are used. They're "Japan," "Last Year," "Government," "Assemblyman," "Reporter," "USA," "Problem," "Article," "South Korea" and so on. The number of occurrence for each word is used as a feature value for the news.

- The number of characters in the title (S_CC)
- The number of words in the title (S_WC)
- The number of characters in the main body (CC)
- The number of words in the main body (WC)
- The number of sentences in the main body (SC)
- The number of paragraphs in the main body (PC)
- The maximum number of characters in paragraphs (MAX_CC)
- The number of pictures (PM)

4.3 Learning intention models

Table 4 summarizes the prediction accuracy of the individual model (10-fold cross validation). For the learning stage, WEKA machine learning library is used (Hall *et al.*, 2009). In the classification, it is formulated as a binary classification problem. The goal is to classify each news article into "summarization" or "no summary or expansion." It has 51 features (preference, topic, time spent, the eight structural features, and 40 words features). The representative classification algorithms have been adopted.

- Bayesian network (Probabilistic approach) (Gelman *et al.*, 2014): It models the probabilistic distribution of data with a graphical structure. In the graph, an edge represents conditional dependency of two variables and the conditional probability table on each node stores parameters of the model. Because the model is based on probability theory, it's strong to handle data with uncertainty.
- Multi-Layered Perceptron (Neural approach): It has been widely used to approximate complex non-linear functions from data. It consists of input, hidden and output layers. In each layer, there are different number of neurons with non-linear functions. Although it is powerful to model non-linear functions, it's hard to interpret the meaning of the classification models.
- Simple Logistics (Linear Logistic Regression): It represents the discrimina-

Table 4. Summary of classification performance on user intentions.

	Bayes Net	MLP	Simple Logistics	Bagging	J48	Average
User 1	65.0	67.5	**72.0**	69.0	61.0	66.9
User 2	82.5	75.5	**84.5**	82.5	78.5	80.7
User 3	82.5	78.5	79.5	**84.0**	77.5	80.4
User 4	90.5	89.5	93.5	**94.5**	92.5	92.1
User 5	70.5	67.5	**71.0**	68.5	67.0	68.9
User 6	79.5	77.0	**81.0**	79.0	77.5	78.8
User 7	79.0	71.5	**80.5**	80.0	75.0	77.2
User 8	84.5	88.5	**89.5**	86.5	87.0	87.2
Average	79.3	76.9	**81.4**	80.5	77.0	79.02

tive function using a weighted sum of input features. Although it's simple, it can generalize well and requires small amount of time to process the data. In addition, it's possible to interpret the weight of each feature as relevance to decision making.

- Bagging (Ensemble approach): In the bagging, it creates a lot of different classifiers trained from randomly sampled training data. Although the size of classification model is a bit large, it can be generalized well on unseen data. The performance is dependent on the selection of base classification algorithm and the size of ensembles.
- J48 (Decision tree): It automatically selects relevant features for classification and sets the most important one as a root node. Each path of the tree can be regarded as a single rule for classification. Because it's based on a tree structure, it's easy to explain the decision mechanism behind the classification.

It shows that the simple logistic regression is the most successful classifier for individual intention modelling. In average, the classifiers could predict the user intentions with accuracy of 81.4%.

In the simple logistic regression, the representation of model is a linear equation of weighted sum of features. Using full training samples, it is possible to build linear equations which show the mechanism behind of the decision making. Table 5 summarizes the equations from the logistic regression training. It's interesting that the equations are very different to each user. For example, a simple constant value is used to classify the samples for user 3 and user 4. However, topic, prefer-

Table 5. The linear equations of features derived from the logistic regression.

	Equations
User 1	−1.76 + [Topic]*0.27 + [South Korea]* −0.13
User 2	−2.15 + [Preference]*0.81
User 3	−0.92
User 4	−0.00
User 5	−0.46 + [Preference]* −0.41 + [S_WC]* −0.07 + [Government]*0.1 + [Reporter]* 0.18 + [Problem]*0.12 + [Market]*0.09 + [New Party]* −0.12 + [Event]*0.2 + [Democratic Party]*0.14 + [Smart Phone]* −0.07 + [Possibility]* −0.24
User 6	−3.98 + [Preference]*0.56 + [Topic]* −0.08 + [Time Spent]*0.02 + [S_CC]*0.01 + [Last_Year]* −0.15 + [Reporter]* −0.13 + [USA]* −0.07 + [Article] *0.11 + [South_Korea]*0.08 + [Market]*0.06 + [News]* −0.43 + [Result]* −0.45 + [Apple]* −0.12 + [Information]*0.32 + [Democratic_Party]*0.17 + [Investigation]*0.15 + [Accident]* −0.05 + [Seoul]* −0.12 + [Word]* −0.17 + [Chairman] *0.8 + [Textbook]* −0.04 + [Conspiracy]* −0.23
User 7	−3.31 + [Preference]* −0.51 + [Topic]*0.28 + [WC]*0.01 + [SC]* −0.02 + [PC]* −0.03 + [Assemblyman]* −0.1 + [Article]* −0.51 + [Case]*0.16 + [Event]*0.2 + [Democratic_Party]* −0.13 + [Chairman]*0.08 + [North_Korea]*0.09 + [World]*0.24
User 8	1.94 + [PC]* −0.04

ence, photo count and the word “South Korea” contributes to the classification of the user 1, user 2 and user 8. Based on the weights assigned for each feature, it is possible to infer the relevant factors for each user on the prediction of “summarization” service access. For user 5, user 6, and user 7, the equations are very complex compared to other users. It includes a lot of words in their decision making. We analyzed the models from different gender but failed to find meaningful conclusion. It seems that the models are highly dependent on individual users' interest.

In the integrated model, all the data from the eight users are used to train single model. The performance of the logistic regression is 73.6%. It's a bit lower than the average of individual models. In this model, the feature on the words inclusion/exclusion is not so useful to predict the user's intention. Because the inclusion/exclusion of words is highly dependent on user's preference, it is better to exclude them in the integrated model learning. The accuracy (75.3%) is increasing if we use only 11 features excluding the 40 words features. The linear equation for the integrated model is as follows.

$$\begin{aligned}&-1.19+[\text{Preference}]*-0.25+[\text{Topic}]*0.05+[\text{S\_CC}]*-0.01\\&+[\text{S\_WC}]*-0.01+[\text{SC}]*-0.01+[\text{PC}]*0.01+[\text{PM}]*0.01\end{aligned}$$

It shows that the user preference and topic of news articles are important to predict user's intention on the summarization. Unexpectedly, the importance of the structural properties (S_CC, S_WC, SC, and PC) is not so significant. It means that the structural property is not the main point to predict user's intention on the summarization. Instead, it is important to know user's preference on the news and the category of the article.

For topic models, we have introduced the same idea in learning models and the 11 features are used to train them. The accuracy is 77.5% for politics, 76.25% for economy, 71.6% for society, 75.3% for world and 71.9% for IT/Science. It shows that the performance is ranging from 71.6% to 77.5%. Table 6 shows the equations used to predict the user intention for topic models. For economy topic, the equation is constant and very simple. For politics, society and world, the user preference is the most important factor for the prediction. However, it's not true for the IT/Science. Instead in the category, the number of photos is the most important

Table 6. The equations derived from linear regression for topic models.

Topics	Equation
Politics	$-1.2+[\text{Preference}]*-0.33+[\text{PC}]*0.03+[\text{PM}]*0.14$
Economy	1.08
Society	$-1.43+[\text{Preference}]*0.24$
World	$0.92+[\text{Preference}]*-0.43+[\text{S\_WC}]*0.07+[\text{SC}]*0.02+$
IT/Science	$-1.32+[\text{Time Spent}]*-0.01+[\text{SC}]*-0.02+[\text{PC}]*0.02+[\text{PM}]*0.05$

feature. For politics, the number of photos has relatively big weight.

5. Conclusions and Future Works

In this paper, we propose to use machine learning on the prediction of user's intention for summarization service. The experimental results show that only 30% of news articles are labeled for "summarization" and 70% are labeled for "no summary" or "expansion." In a straightforward guess, the summarization might be required for the news with long length. However, the study shows that the structural property is not always useful to predict the user intention. Because the cognitive process is not explained simply with the length of article, it is necessary to learn models from data. In the experiments, we tested individual, integrated, and topical models. The comparison showed that the linear regression approach is the best classifier for the individual intention models. Also, it allows designers to understand the decision mechanism easily from a simple linear equation. In the integrated and topical models, the exclusion of word information is useful to get high accuracy.

The contribution of this paper is to introduce the problem of prediction on user's intention for news article manipulation (summarization, no summary, and expansion). Still, there is little information on the human's cognitive functions for the task. The machine learning approach shows that the prediction on the user's intention on news manipulation could be predicted with accuracy over 80%. Also, the analysis of the successful models could give insight on the human cognition for data processing (summarization and expansion). We demonstrated the possibility of our approach from the user survey (80 participants) and user testing (8 participants).

Although it could predict with accuracy over 80%, there are variation on the performance over users. For the most successful users, the accuracy is over 90% but the worst case shows 70%. It is necessary to understand the reason of the variation and develop a method to reduce the gap. Also, there are about 10% news articles labeled by "expansion." It's interesting to provide a mechanism to predict user's intention on "expansion" and provide automatic expansion service. Unlike the summarization, there is little work on the data expansion.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (2013 R1A2A2A01016589, 2010-0018950).

References

Abel, F., Q. Gao, G.-J. Houben and K. Tao (2011) "Analyzing user modeling on twitter for personalized news recommendations," *User Modeling, Adaptation and Personalization (Lecture Notes in Computer Science)* 6787: 1–12.

Aliguliyev, R. M. (2009) "A new sentence similarity measure and sentence based extractive technique for automatic text summarization," *Expert Systems with Applications* 36.4: 7764–7772.

Antiqueira, L., O. N. Oliveira Jr., L. Costa and M. Nunes (2009) "A complex network approach to text summarization," *Information Science* 179.5: 584–599.

Barzilay, R. and M. Elhadad (1997) "Using lexical chain for text summarization," in *Proceedings of the ACL Workshop on Intelligence Scalable Text Summarization*, pp. 10–17.

Cho, S. B., K. J. Kim, K. S. Hwang and I. J. Song (2007) "AniDiary: Summarizing user's daily life in cartoon-style diary with Bayesian networks," in *IEEE Pervasive Computing / IEEE Computer Society [and] IEEE Communications Society*, pp. 66–75.

Diaz, A. and P. Gervas (2007) "User-model based personalized summarization," *Information Processing & Management* 43.6: 1715–1734.

Endres-Niggemeyer, B. and E. Wansorra (2004) "Making cognitive summarization agents work in a real-world domain," in *Proceedings of the 1st International Workshop on Natural Language Understanding and Cognitive Science*, pp. 86–96.

Erkan, G. and D. R. Radev (2004) "LexRank: Graph-based lexical centrality as salience in text summarization," *Journal of Artificial Intelligence Research* 22: 457–479.

Fattah, M. A. and F. Ren (2009) "GA, MR, FFNN, PNN and GMM based models for automatic text summarization," *Computer Speech & Language* 23.1: 126–144.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari and D. B. Rubin (2014) *Bayesian Data Analysis*, 3rd ed., CRC Press.

Gong, Y. and X. Liu (2001) "Generic text summarization using relevance measure and latent semantic analysis," in *Proceedings of the 24th Annual International ACM SIGIR Conf. on Research and Developement in Information Retrival*, pp. 19–25.

Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. H. Witten (2009) "The WEKA data mining software: An update," *SIGKDD Explorations* 11.1: 10–18.

Horvitz, E., S. Dumais and P. Koch (2004), "Learning predictive models of memory landmarks," *26th Annual Meeting of the Cognitive Science Society*.

Hosmer, D. W. and S. Lemeshow (2013) *Applied Logistic Regression*, 3rd ed., Wiley.

Kim, K. J., M. C. Jung and S. B. Cho (2007) "KeyGraph-based chance discovery for mobile contents management system," *International Journal of Knowledge-based Intelligent Engineering System* 11: 313–320.

Lee, J. E., H. S. Park, K. J. Kim and J. C. No (2013) "Learning to predict the need of summarization on news articles," *Procedia Computer Science* 24: 274–279.

Li, L., W. Chu, J. Langford and R. E. Schapire (2010) "A contextual-bandit approach to personalized news article recommendation," in *Proceedings of the 19th International Conference on World Wide Web*, pp. 661–670.

Lin, C. and E. Hovy (2000) "The automated acquisition of topic signatures for text summarization," in *Proceedings of the 18th Conference on Computational Linguistics 1*, pp. 495–501.

Liu, J., P. Dolan and E. R. Pedersen (2010) "Personalized news recommendation based on click behavior," in *Proceedings of the 15th International Conference on Intelligent User Interfaces*, pp. 31–40.

Nenkova, A., A. Siddharthan and K. Mckeown (2005) "Automatically learning cognitive status for multi-document summarization of newswire," in *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Lan-*

guage Processing, pp. 241–248.
Nenkova, A. and K. McKeown (2012) "A survey of text summarization techniques," *Mining Text Data*, pp. 43–76.
Quinlan, J. R. (1996) "Bagging, boosting, and C4. 5," in *AAAI Proceedings of the Thirteenth National Conference on Artificial Intelligence 1*: 725–730.
Reichert, D. and D. Kaufman B.Bloxham H. Chase and N. Elhadad (2010) "Cognitive analysis of the summarization of longitudinal patient records," in *American Medical Informatics Association Annual Symposium Proceedings*, pp. 667–671.
Ruck, D. W., S. K. Rogers, M. Kabrisky, M. E. Oxley and B. W. Suter (1990) "The multilayer perceptron as an approximation to a Bayes optimal discriminant function," *IEEE Transactions on Neural Networks* 1.4: 296–298.
Salton, G., A. Singhal, M. Mitra and C. Buckley (1997) "Automatic text structuring and summarization," *Information Processing & Management* 33.2: 193–207.