NAÏVE BAYES AND BLACK BOX TESTING IMPLEMENTATION ON SENTIMENT ANALYSIS OF ALOE VERA PRODUCT REVIEWS

Putri Ambarwati

Master of Business Information Systems
Technology and Engineering Master Program
Gunadarma University
www.gunadarma.ac.id
putriambar16@gmail.com

Abstract— Aloe vera soothing gel is one of the best-selling products and the most widely reviewed on the Althea Korea website. This product has been reviewed by 1,448 users on the Althea website. The result of the research can be used to minimize mistakes in product purchases. Besides, through a review of a product, the company can analyze the level of customer satisfaction and can be a suggestion for improvements in the future. Therefore, a system is needed to analyze the sentiment towards aloe vera soothing gel to determine the review as a positive or negative sentiment. The method used in this research is the Naïve Bayes method and uses the classification carried out by linguists as a reference for determining positive and negative sentiment. There are two tests carried out in this research, namely confusion matrix testing and black-box testing. The result of the confusion matrix test found an accuracy of 94.62% and the result of black-box testing showed that the output produced was by the application functionality.

Keywords: Black-Box Testing; Naïve Bayes; Product Review; Sentiment Analysis

Abstrak—*Produk aloevera soothing gel merupakan salah satu produk dengan penjualan terbaik di website Althea Korea. Produk ini sudah diulas oleh 1.448 pengguna di website Althea. Hasil penelitian dapat digunakan untuk meminimalkan kesalahan dalam pembelian produk. Selain itu, melalui ulasan produk, perusahaan dapat menganalisis tingkat kepuasan pelanggan dan dapat menjadi saran untuk perbaikan di masa depan. Oleh karena itu, diperlukan suatu sistem untuk menganalisis sentimen terhadap produk aloevera soothing gel untuk menentukan ulasan tersebut merupakan ulasan yang positif atau negatif. Metode yang digunakan dalam penulisan ini adalah metode Naïve Bayes dan menggunakan klasifikasi yang dilakukan oleh ahli bahasa sebagai acuan untuk menentukan kalimat positif dan negatif. Ada dua pengujian yang dilakukan dalam penelitian ini, yaitu pengujian confusion Matrix dan pengujian black box. Hasil dari pengujian confusion Matrix didapatkan akurasi sebesar 94,62% dan hasil pengujian black box testing menunjukkan bahwa output yang dihasilkan sudah sesuai dengan fungsionalitas aplikasi.*

Kata Kunci: *Analisis Sentimen; Naïve Bayes; Ulasan Produk; Pengujian Black Box.*

INTRODUCTION

Women's needs for skincare products are increasing. Based on a survey conducted by Bizteka-CCI in 2015, the national cosmetics market is estimated to grow by 8.3% with a value reaching Rp. 13.9 trillion, an increase compared to the previous year (2014) which amounted to Rp. 12.8 trillion. During the 2010-2015 period, the national cosmetic industry market increased by an average of 9.67% per year (Rohmatun & Dewi, 2017).

Not only used for aesthetic functions but now cosmetics also play a role in healing and skincare. Although not a primary need, cosmetics are one of the products that are used routinely and continuously by humans. Therefore, the safety of cosmetics from hazardous substances needs to be considered (Awalia, 2018). To avoid products that can have negative effects on the skin, looking for as much review information as possible about a beauty product is one way to minimize mistakes in purchasing a product. Currently, information about beauty products is very easy to find. The abundance of information on beauty products cannot be separated from their services who share their experiences after using beauty products through social media, personal blogs, or websites that specifically have features to review beauty products. (Pujadayanti et al., 2018). One website that has a feature to review beauty products is Althea Korea.

Althea is an e-commerce for Korean beauty products that was launched in July 2015 with branches in Malaysia, Singapore, Indonesia, the Philippines, Thailand, and the United States. (Zatayumni et al., 2018). Althea offers products that are priced similar to the original prices in

P-ISSN: 1978-2136 | E-ISSN: 2527-676X | Naïve Bayes and Black...
Techno Nusa Mandiri : Journal of Computing and Information Technology
As an Accredited Journal Rank 4 based on **Dirjen Risbang SK No. 85/M/KPT/2020**

Korea so that consumers around the world can enjoy authentic Korean products at affordable prices, the same as consumers in Korea. (Pawestri et al., 2019). In 2017, Althea Korea expanded its business by opening worldwide shipping services, which means all of his current achievements have been achieved in just two years (Sekaringtias & Kusumawati, 2017). One of the best selling and most reviewed products on the Althea Korea website is the aloe vera gel product from the Nature Republic brand which has 1,448 reviews (https://althea.kr/ accessed on 27 January 2019 at 10.00 a.m).

Following the official Instagram of Nature Republic Indonesia (@ Naturerepublic.id), this aloe vera gel product is claimed to moisturize dry skin and redness caused by the sun. Aloe vera gel can moisturize the skin because it contains a lot of water. Aloe vera gel has a high water content of 94.83% (Khoirini, 2018). Besides, Aloe vera gel can cool and has many benefits, namely as a natural moisturizer, face freshener, prevent wrinkles, help regenerate skin cells, brighten skin, treat acne, remove scars, treat hair and reduce stretch marks. (Noormindhawati & ReMediaService, 2016). According to Dream.co.id, unfortunately, not everyone is compatible with aloe vera gel. (Male, 2019). Based on a review of the aloe vera gel product from the Nature Republic brand on the Althea Korea app, some people say the product makes their skin spotty and doesn't help moisturize their skin.

Based on the description above, research was carried out on the implementation of Naïve Bayes and black box testing on the sentiment analysis of Aloe vera gel product reviews from the Nature Republic brand on the website https://althea.kr/. This study aims to minimize mistakes in product purchases. Besides, through product reviews, the company can analyze the level of customer satisfaction and can provide suggestions for future improvements.

Research on sentiment analysis has been done before. Among them is research conducted by Kristiyanti (Kristiyanti, 2015) which is about sentiment analysis that reviews cosmetic products using a support vector machine and a particle swarm optimization algorithm with an accuracy rate of 89% (Kristiyanti, 2015). Research conducted by Rosi (Rosi et al., 2018) on prediction ratings in beauty product reviews uses the naïve Bayes method and the categorical proportional difference (CPD) with an accuracy of 87%. (Rosi et al., 2018).

This paper uses the Naïve Bayes method which is intended to process and classify opinions so that the classification of these opinions will be known as positive or negative. The Naïve Bayes method has several advantages, one of them is an efficient computation method and also has a good predictive performance. (Somantri & Khambali, 2017). One of the tests in this paper uses black-box testing. The advantages of black-box testing are that the number of test cases can be reduced to achieve reasonable testing and test cases can indicate the presence or absence of an error. (Ardana, 2019).

This paper also uses the R Studio software as a tool used to process data. R Studio is one of the integrated development environments (IDE) for R. R is a well-known environment for computational statistics and graphs used in statistical and data analysis tasks. (Salim et al., 2018). R is a programming language that uses command-line scripts for graphical and statistical analysis and representation and ultimately generates reports. R is free, open-source, robust, and highly extensible for data analysis (Rajput et al., 2019).

MATERIALS AND METHODS

This paper uses the Naïve Bayes method. The following are the stages of sentiment analysis using the naïve Bayes method.

1. Data Source Analysis

The data of the review were taken from the website Althea Korea (https://althea.kr/). Data uses English text. Data is input manually to Microsoft Excel. After the data were gathered, they were then validated. The data were labeled with the positive or negative class manually. There are 313 reviews gathered for the research. They consist of 50 negative reviews and 263 positive reviews. From those data, 220 data were used as training data, and 93 data were used as testing data. The research used a supervised learning approach so that the existing review data were divided into training data and testing data.

2. Data Preprocessing

The gathered review data were not well structured yet, therefore, some preprocessing steps were required to make them well structured. Structured data simplify the classification process. Preprocessing was conducted through some steps, including case folding, tokenization, stopwords removal, and stemming.

- Case folding is a font conversion step, converting all letters to lowercase letters.
- Tokenization is used to remove redundant numbers, symbols, and spaces.
- Stop-words are used to eliminate meaningless words.
- Stemming is used to reduce various grammatical forms/tenses from words such as nouns, adjectives, verbs, verbs, etc.

P-ISSN: 1978-2136 | E-ISSN: 2527-676X | Naïve Bayes and Black ...
Techno Nusa Mandiri : Journal of Computing and Information Technology
As an Accredited Journal Rank 4 based on **Dirjen Risbang SK No. 85/M/KPT/2020**

3. Sentiment classification by a linguist

The sentiment classification process for data was carried out manually by an English teacher who had a bachelor's degree and a master's degree in English literature. Table 1. Shows some of the classifications of sentiments carried out by a linguist.

Table 1. Data classification by a linguist

Date	Name	Reviews	Results
26/1/2019	Ku B.	Affordable and easy to absorb	Positive
24/1/2019	Josephine B.	I LOVE IT. The best moisturizer!	Positive
23/1/2019	Aimee C.	Favorite moisturizer. I love this product. It's light and non-greasy. It's perfect for combination skin.	Positive
21/1/2019	Maria	It lightens my acne marks	Negative

Source : (Ambarwati, 2019)

4. Sentiment classification using the Naïve Bayes method

The sentiment classification stage using the Naïve Bayes algorithm is used to determine which reviews from users are positive or negative opinions. Figure 1 is a flowchart of the sentiment classification stage using the Naïve Bayes method.
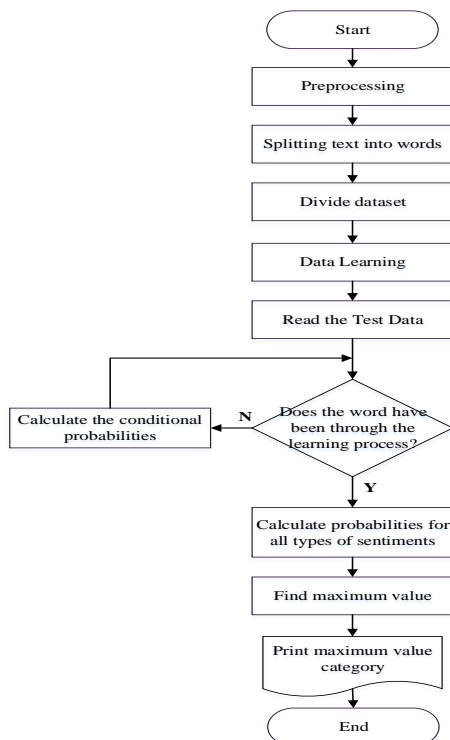
Source : (Ambarwati, 2019)

Figure 1. Flowchart of the Naïve Bayes method

RESULTS AND DISCUSSION

The discussion in this paper discusses the ways to get results of conducting sentiment analysis. Accompanied by two tests, namely confusion matrix, and black-box testing.

1. Data Pre-processing

Data Preprocessing is a proven method for solving incomplete or inconsistent data problems. Cleaning the data in RStudio required several packages including the "tm" package, the "Stopwords" package, and the "SnowballC" package. Table 2 shows an example of the results of the preprocessing process of data.

Table 2. Pre-processing

	Before Pre-processing	After Pre-processing
Case folding	I LOVE IT. The best moisturizer!	i love it. the best moisturizer!
Tokenization	i love it. the best moisturizer!	i love it the best moisturizer
Stopwords	i love it the best moisturizer	love the best moisturizer
Stemming	love the best moisturizer	love the best moisture

Source : (Ambarwati, 2019)

2. Document Term Matrix

After the preprocessing process, the data will be converted into a corpus and then processed in the form of a matrix using the Document Term Matrix (DTM). The term matrix document is used to separate text documents into words.

3. Divided Datasets

At this stage, divide the dataset into two subsets as the training dataset or training data to train the Naïve Bayes model and the test dataset or test data to validate and make predictions. From these data, 220 data are used as training data taken from 70% of all data, and 93 data are used as test data taken from 30% of all data.

4. Words Frequency

The document terms matrix lists all occurrences of the word in the document corpus represented by columns. If a word appears in the document, then the matrix entry matching the row and column is 1 (one), and otherwise 0 (zero). For every word that appears in the matrix, all words can't be used for classification. So to reduce the

number of words is to eliminate words that have low frequency using the findfreqterm () function in the programming language R.

5. Training and evaluating the naïve Bayes model on data.

At this stage, the implementation of Naïve Bayes is implemented using the e1071 package included in the R-package. Then create a Naïve Bayes model using previous training data and create predictive data using test data based on the Naïve Bayes model. Table 3. is a comparison of data classifications performed by linguists and Naïve Bayes using test data.

Table 3. Comparison of classification

Test Data	Sentiment Classification	
	Linguist	Naïve Bayes
Recommended. Good to use	Positive	Positive
This is a gem! Super love this product. It can be used in different ways to help you in different face issues	Positive	Positive
Plus the tube container is unhygienic and inconvenient to use and It's not moisturizing my skin	Negative	Positive
Very authentic for now my face is starting to clear up so thank you so much althea!	Negative	Positive

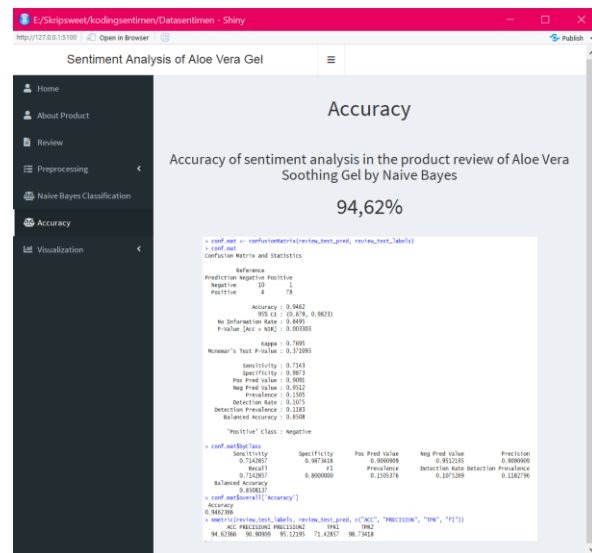Source : (Ambarwati, 2019)

6. Interface Implementation

In this implementation phase, it displays the system output that has been made. This implementation displays the results of sentiment analysis in the form of visualizations using the Shiny package and the Dashboard Shiny package from RStudio. There are seven main interface menus created, namely the home menu, product menu, review menu, preprocessing menu consisting of four submenus, Naïve Bayes classification menu, accuracy menu, and visualization menu which consists of three submenus. Some of the menu interfaces can be seen in Figure 2, Figure 3, and Figure 4.
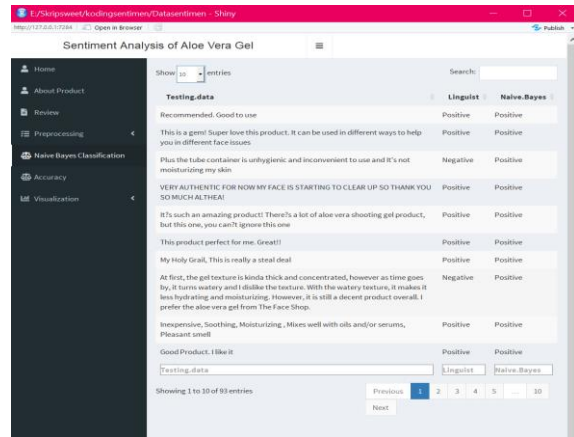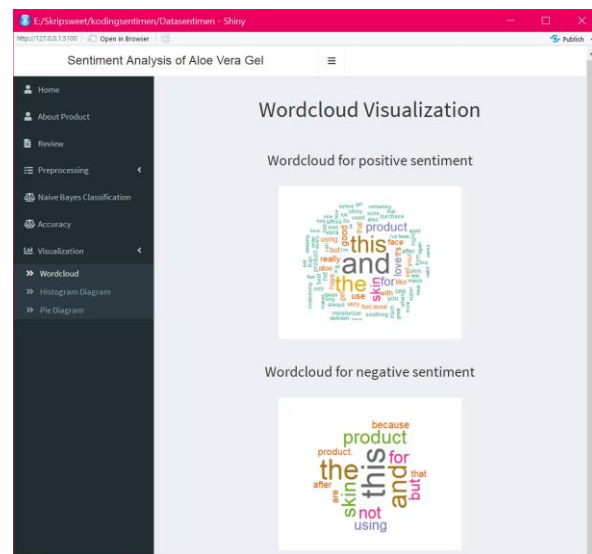

Source : (Ambarwati, 2019)
Figure 2. The Interface of Naïve Bayes Classification Results

Source : (Ambarwati, 2019)
Figure 3. The Interface of Accuracy Results

Source : (Ambarwati, 2019)
Figure 4. The Interface of Visualization Results

P-ISSN: 1978-2136 | E-ISSN: 2527-676X | Naïve Bayes and Black ...
Techno Nusa Mandiri : Journal of Computing and Information Technology
As an Accredited Journal Rank 4 based on Dirjen Risbang SK No. 85/M/KPT/2020

7. Testing

After the syntax compiles successfully, several tests must be run to ensure the application is functioning properly and to analyze the performance of the application on the system on which it is located. Testing is done in two ways, namely testing with confusion matrix to measure algorithm accuracy and black box testing to test application functionality.

a. Confusion Matrix

Testing has been done manually by calculating accuracy using the confusion matrix theory formula. The data used in this test is test data which amounts to 93 data. Testing is done by comparing the results of the Naïve Bayes classification with the data classification performed by linguists. Then the comparison results are summarized into four parts including true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) as shown in Table 4.

Table 4. Confusion Matrix

Amount of Data: 93	Naïve Bayes Classification		Total
Classification by Linguist	Positive	Negative	
Positive	True Positives (78)	False Negatives (1)	9
Negative	False Positives (4)	True Negatives (10)	42
Total	14	79	

Source : (Ambarwati, 2019)

The following is the result of the accuracy calculation.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Accuracy = \frac{10 + 78}{10 + 78 + 4 + 78}$$

$$Accuracy = 0{,}9462365591$$

$$Accuracy = 94{,}62\%$$

b. Black-box Testing

Black box testing to determine whether all software functions run properly by the functional requirements that have been designed. The black box testing method allows software engineers to derive a set of input conditions that fully utilize all functional requirements for a program. The results of the study using black-box testing, it can be concluded that all system interface components are

by their function. Table 5 is the black box test results. Status: S for success and F for failure.

Table 5. The results of black-box testing

Form	Scenario	Results	Status
Home Menu	Access the home menu	Displays the home menu containing the author's identity	S
About Product Menu	Access the About Product Menu	Displays about product menu that contains information about the product being analyzed	S
Reviews Menu	Access the Reviews Menu	Displays a reviews menu that contains all raw data in the tabular form taken from the Althea website.	S
Pre-proces-sing Menu	Access the Pre-proces-sing Menu	Displays all processed data in preprocessing form which consists of 4 parts, namely Case folding, Tokenization, Stopwords, and Stemming.	S
Naïve Bayes Menu	Access the Naïve Bayes Menu	Displays test data and sentiment classification results.	S
Accuracy Menu	Access the Accuracy Menu	Displays the results of research accuracy.	S
Visualiza tion Menu	Access the Visualizat ion Menu	Displays data in the form of a word cloud, histogram, and pie diagram.	S

Source : (Ambarwati, 2019)

CONCLUSION

In this paper, the implementation of Naïve Bayes and black box in the sentiment analysis of aloe vera product reviews have been completed. This paper uses a classification conducted by linguists as a reference to determine positive and negative sentiments. The data in this study were taken from the Althea Korea website as many as 313 data. The data was divided into two parts including training data for 220 data and test data for 93 data. This research has also been tested with two tests, namely confusion matrix testing to measure the performance and accuracy of the algorithm and black box testing to test application functionality. The results of this study obtained an accuracy of 94.62% of the 93 data tested. Of the 93 data, 87

P-ISSN: 1978-2136 | E-ISSN: 2527-676X | Naïve Bayes and Black...
Techno Nusa Mandiri : Journal of Computing and Information Technology
As an Accredited Journal Rank 4 based on Dirjen Risbang SK No. 85/M/KPT/2020

data from the naïve Bayes classification have results that are suitable by the classification carried out by linguists. Some of the causes of the accuracy value being less than 100% due to lack of training data, doesn't use negation handling, and some words being processed incorrectly by the stemming process, such as the word easy to easi, the word morning to be morn, etc. The results of the black box test show that the output produced is by the functionality of the application.

REFERENCES

Ambarwati, P. (2019). *Sentiment Analysis Towards Reviews of Aloe Vera Soothing Gel using Naive Bayes Method in English Text.* Universitas Gunadarma.

Ardana, I. M. S. (2019). Pengujian Software Menggunakan Metode Boundary Value Analysis dan Decision Table Testing. *Jurnal Teknologi Informasi ESIT*, *14*(11), 40–47.

Awalia, F. T. M. (2018). PERILAKU KONSUMSI KOSMETIK HALAL DALAM MEMBANGUN PERADABAN. *International Conference of Afro-Asian University Forum (AAUF) on the Role of the Afro-Asian Universities in Building Civilizations*, 1183–1201.

Khoirini, F. (2018). Gel Lidah Buaya dalam Mengurangi Pruritus. *Jnph*, *6*(2), 25–34.

Kristiyanti, D. A. (2015). Analisis Sentien Review Produk Kosmetik Melalui Komparasi Feature Selection. *Konferensi Nasional Ilmu Pengetahuan Dan Teknologi (KNIT)*, 69–78.

Male, C. A. (2019). *Tidak Cocok Pakai Aloe Vera Gel? Ini Sebabnya...* https://www.dream.co.id/beauty/tidak-cocok-pakai-aloe-vera-gel-ini-sebabnya-190913h.html.

Noormindhawati, L., & ReMediaService. (2016). *Raja Obat Alami Aloe Vera Khasiat A-Z untuk Kesehatan dan Kecantikan* (Maya (ed.)). Rapha Publishing.

Pawestri, R. H., Az-Zahra, H. M., & Rusydi, A. N. (2019). Evaluasi Usability Aplikasi Mobile menggunakan Usability Testing dan System Usability Scale (SUS) (Studi Kasus: SOCO, Althea dan Sephora). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK)*, *3*(10), 9883–9891.

Pujadayanti, I., Fauzi, M. A., & Sari, Y. A. (2018). Prediksi Rating Otomatis pada Ulasan Produk Kecantikan dengan Metode Naïve Bayes dan N-gram. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK)*, *2*(11), 4421–4427.

Rajput, D. S., Thakur, R. S., & Basha, S. M. (2019). *Sentiment Analysis and Knowledge Discovery in Contemporary Business*. IGI Global.

Rohmatun, K. I., & Dewi, C. K. (2017). Pengaruh Pengetahuan Dan Religiusitas Terhadap Niat Beli Pada Kosmetik Halal Melalui Sikap. *Journal Ecodemica*, *1*(1), 27–35.

Rosi, F., Fauzi, M. A., & Perdana, R. S. (2018). Prediksi Rating Pada Review Produk Kecantikan Menggunakan Metode Naïve Bayes dan Categorical Proportional Difference (CPD). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *2*(5), 1991–1997.

Salim, N. R., Gopal, K., & Ayub, A. F. M. (2018). Experiential Statistics Learning with RStudio: Study on Students' Engagement. *Journal of Physics: Conference Series*, *1132*(1). https://doi.org/10.1088/1742-6596/1132/1/012039.

Sekaringtias, K., & Kusumawati, A. (2017). PERAN PERMISSION EMAIL MARKETING DALAM MEMEDIASI PENGARUH TRUST TERHADAP PURCHASE INTENTION (Studi pada Konsumen Althea Korea Di Indonesia dan Malaysia). *Jurnal Administrasi Bisnis*, *50*(6), 174–182.

Somantri, O., & Khambali, M. (2017). Feature Selection Klasifikasi Kategori Cerita Pendek Menggunakan Naïve Bayes dan Algoritme Genetika. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi (JNTETI)*, *6*(3), 301–306. https://doi.org/10.22146/jnteti.v6i3.332.

Zatayumni, F., Rachmadi, A., & Herlambang, A. D. (2018). Fakultas Ilmu Komputer Universitas Brawijaya 2937 Analisis Kualitas Layanan Website Althea Indonesia Terhadap Kepuasan dan Loyalitas Pelanggan Menggunakan E-S-Qual dan E-RecS-Qual. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK)*, *2*(9), 2937–2947.