

A Statistical View to Study the Aphorisms in Nahj al-Balaghah

Yu Tian

College of Literature, Northwest University, Xi'an City, Shaanxi Province, China

Kim-Hung Pho

*Fractional Calculus, Optimization and Algebra Research Group, Faculty of Mathematics
and Statistics, Ton Duc Thang University, Ho Chi Minh City, Vietnam*

Mohammad Reza Mahmoudi

Department of Statistics, Faculty of Science, Fasa University, Fasa, Iran

Abstract. Besides the Holly Quran, Nahj al-Balagha is the main source of literature in Arabian nations, specially for Shia as one of the main branches of the Muslims. Along with literary brilliance, the text of this book covers the broad topics. This research deals with the application of the statistical text and data analysis to extract knowledge from the aphorisms in Nahj al-Balaghah. First, we classify these aphorisms in seven topics. Then, the count of the aphorisms in each category is computed. Finally, the counts of the aphorisms of the categories are compared using the chi-square test and the cluster analysis.

Keywords: Aphorisms, Nahj al-Balagha, Statistics, Data Analysis, Text Mining.

1. Introduction

It is taken for granted that one of the famous and important Islamic books is Nahj al-Balagha. It would be worth mentioning that this book belongs to Ali, cousin and son-in-law of Muhammad and first imam of Shia and so forth caliph of Sunni Muslims. Ali's words have greatly influenced in Arabic language and literature such as tajol al- Arus and lisan al-Arab.

Nahj al-Balagha is the main source for recognizing the intellectual foundations of Shia as one of reflective movements of Muslims. Along with literary brilliance, this book covers the broad topics of monotheism, social, moral, and political, so that it can be called the "book of life". In regard to Nahj al Balagha, it is a collection of 241 sermons, 79 letters and 480 Aphorisms.

One of the characteristics of this book is its numerous translations, in such a way that it can be said that in the domain of Islamic civilization, after the Holy Qur'an, no book has had as many translations as Nahj al-Balaghah. It is noteworthy that this book collected by Sharif Razi, a Shia scholar who live in the tenth century.

2. ALĪ B. ABĪ ṬĀLEB and Aphorism

Ali is known as the first Imam of all Shia sects and the fourth Muslim Caliphe. He was born in the holiest place in Islam, Ka'bah, in Mecca according to some writers, he was the first man to accept Islam and he is the first Muslim. He governs as a caliph, less than five years from the age of 656 to 661. He has a prominent position among a large group of Muslims. Ali's sermons and aphorisms are regarded as the main reference of philosophical knowledge or divine wisdom, and he is considered the first person to use Arabic terms for expressing philosophical opinions. A large number of Ali's short words have become parts of Islamic general culture and

are referred to as proverbs and utterances in everyday life. Also, these words are the basis of literary works and come in poems in various languages.

An aphorism denotes opinion or makes a statement of wisdom and it is a brief phrase expresses some truth or principle concept. Aphorism-Ali's words contain various subjects, and these sentences are widely seen in the language of the people and the works of the poets. The concepts of these sentences can be divided into the following: Political, Moral, Social, Economical, Religious beliefs, World beliefs and Jurisprudence.

This research deals with the aphorisms in Nahj al-Balaghah. First, we classify these aphorisms in seven categories. Then, the count of the aphorisms in each category is computed. Finally, the counts of the aphorisms of the categories are compared using chi-square test and cluster analysis.

3. Methodology

This section is devoted to discuss about the data collection, data analysis and data analysis techniques. The details of the collected dataset are expressed in the First Subsection. Subsection Two summarizes some studies that applied different data analysis techniques. In Subsection Three, the data analysis techniques that are used to investigate the considered dataset are given.

3.1. Data Collection

In this research, all aphorisms in Nahj al-Balaghah are covered. We classify these aphorisms in seven categories: Political, Moral, Social, Economical, Religious beliefs, World beliefs and Jurisprudence.

3.2. Data Analysis

Text mining analysis is used to extract knowledge from text datasets. Mahmoudi et al. (2018) studied the alternatives of the word “Love” in Divan of Moulana. First, these words were determined and their counts were computed. Then, they classified the words in eleven categories. Finally, the counts of the words and the categories were compared using chi-square test and cluster analysis. Mahmoudi and Abbasalizadeh (2018a) studied the alternatives of the word “Love” in Saadi’s poetry. First, these words were determined and their counts were computed. Then, they classified the words in twelve categories. Finally, the counts of the words and the categories were compared using chi-square test and cluster analysis. Mahmoudi and Abbasalizadeh (2018b) studied the Divan of Khaghani. They classified the lines into three categories and compared the frequency of different categories for the lines and odes. Also, they investigated the existence of trend in the consecutive lines. Mahmoudi and Abbasalizadeh (2018c) studied the similarities between different orders of Quran’s revelation by using regression analysis and hierarchical clustering method. Mahmoudi and Abbasalizadeh (2018d) applied different text analysis methods to study the Divan of Khaghani. Mahmoudi and Abbasalizadeh (2018e) studied the God’s traits in Quran. They investigated that which group of God’s traits (beauty or divine grace) is more repeated and also what difference there is between the Meccan and Medinan suras in the description of God.

3.3. Statistical Methods

The collected dataset are analyzed using the SPSS 25 and R 3.4.1. First, the descriptive statistics about the counts for the categories of aphorisms in Nahj al-Balaghah are reported. Then, the chi-square test is employed to compare the counts of the categories. Finally, these words and categories are clustered by applying the K-Means clustering technique.

4. Results

This section deals with the results of data analysis. The descriptive statistics about the counts for the categories of aphorisms in Nahj al-Balaghah are reported in Subsection One. The results of chi-square tests and K-Means clustering technique are respectively reported in Subsections Two and Three.

4.1. Descriptive Statistics

Table 1 summarizes the counts and relative counts (percent) of the categories of aphorisms in Nahj al-Balaghah.

Table 1: The counts and relative counts (percent) of the categories of aphorisms in Nahj al-Balaghah

Category		Count	Percent
	World beliefs	210	36.2
	Religious beliefs	132	22.8
	Moral	109	18.8
	Social	58	10.0
	Economical	33	5.7
	Political	22	3.8
	Jurisprudence	16	2.8
	Total	480	100.0

Table 1 indicates that the category “World beliefs” with count=210 (36.2%) and the category “Jurisprudence” with count=16 (2.8%) have the maximum and minimum applications in Nahj al-Balaghah, respectively.

4.2. Comparison between the alternatives

This subsection reports the results of chi-square test to compare the counts of the categories of aphorisms in Nahj al-Balaghah. As can be seen in Table 2, there are significant differences between the counts of the considered categories ($p < 0.001$).

Table 2: The results of chi-square test to compare the counts of the categories of aphorisms in Nahj al-Balaghah

	Chi-Square Test Statistic	P-Value (p)
Category	368.6	<0.001

4.3. Clustering Analysis

In this subsection, we clustered the categories of aphorisms in Nahj al-Balaghah into three groups (high count, medium count, and low count) using their counts. As can be seen in Table 3, the category “World beliefs” the most applications in Nahj al-Balaghah.

Table 3: The results of K-Means clustering to cluster the categories of aphorisms in Nahj al-Balaghah

	High Count	Medium Count	Low Count
Category	World beliefs	Religious beliefs Moral	Social Economical Political Jurisprudence

5. Discussion

Nahj al-Balagha is the main source for recognizing the intellectual foundations of shia as one of reflective movements of Muslims. Along with literary brilliance, this book covers the broad topics of monotheism, social, moral, and political, so that it can be called the "book of life".

This research was devoted to study the aphorisms in Nahj al-Balaghah. First, we classified these aphorisms in seven categories: Political, Moral, Social, Economical, Religious beliefs, World belief and Jurisprudence. Then, the count of the aphorisms in each category was computed. The results indicated that there are significant differences between the counts of the considered categories ($p < 0.001$). Finally, the counts of the aphorisms of the categories were compared using chi-square test and cluster analysis. The clustering results showed that the category “World beliefs” has the most applications in Nahj al-Balagha.

References

Mahmoudi, M. R., Abbasalizadeh, A. (2018a) ‘On comparing and clustering the alternatives of love in Saadi’s lyric poems (Ghazals)’, *Digital Scholarship in the Humanities*. DOI:10.1093/llc/fqy024. In Press.

Mahmoudi, M. R., Abbasalizadeh, A. (2018b) ‘Statistical analysis about the order of Quran’s revelation’, *Digital Scholarship in the Humanities*. DOI:10.1093/llc/fqy030. In Press.

Mahmoudi, M. R., Abbasalizadeh, A. (2018c) ‘How Statistics and Text Mining Can be Applied to Literary Studies?’, *Digital Scholarship in the Humanities*. In Press.

Mahmoudi, M. R., Abbasalizadeh, A. (2018d) ‘Analysis of Mystical Concepts in Khaghani’s Divan’, *Digital Scholarship in the Humanities*. In Press.

Mahmoudi, M. R., Abbasalizadeh, A. (2018e) ‘Statistical Analysis about the God’s Traits in Quran’, *Digital Scholarship in the Humanities*. In Press.

Mahmoudi, MR, Abbasalizadeh, A and Rahmati, M. (2018) *'An Statistical Approach to Investigate the Alternatives of Love in Moulana's Divan'*, *International Journal of Business Intelligence and Data Mining*, In Press.

Samadianfard, Saeed, et al. "Wind speed prediction using a hybrid model of the multi-layer perceptron and whale optimization algorithm." *Energy Reports* 6 (2020): 1147-1159.

Taherei Ghazvinei, Pezhman, et al. "Sugarcane growth prediction based on meteorological parameters using extreme learning machine and artificial neural network." *Engineering Applications of Computational Fluid Mechanics* 12.1 (2018): 738-749.

Qasem, Sultan Noman, et al. "Estimating daily dew point temperature using machine learning algorithms." *Water* 11.3 (2019): 582.

Mosavi, Amir, and Atieh Vaezipour. "Reactive search optimization; application to multiobjective optimization problems." *Applied Mathematics* 3.10A (2012): 1572-1582.

Shabani, Sevda, et al. "Modeling pan evaporation using Gaussian process regression K-nearest neighbors random forest and support vector machines; comparative analysis." *Atmosphere* 11.1 (2020): 66.

Ghalandari, Mohammad, et al. "Aeromechanical optimization of first row compressor test stand blades using a hybrid machine learning model of genetic algorithm, artificial neural networks and design of experiments." *Engineering Applications of Computational Fluid Mechanics* 13.1 (2019): 892-904.

Mosavi, Amir. "Multiple criteria decision-making preprocessing using data mining tools." *arXiv preprint arXiv:1004.3258* (2010).

Karballaezadeh, Nader, et al. "Prediction of remaining service life of pavement using an optimized support vector machine (case study of Semnan–Firuzkuh road)." *Engineering Applications of Computational Fluid Mechanics* 13.1 (2019): 188-198.

Asadi, Esmaeil, et al. "Groundwater quality assessment for sustainable drinking and irrigation." Sustainability 12.1 (2019): 177.

Mosavi, Amir, and Abdullah Bahmani. "Energy consumption prediction using machine learning; a review." (2019).

Dineva, Adrienn, et al. "Review of soft computing models in design and control of rotating electrical machines." Energies 12.6 (2019): 1049.

Mosavi, Amir, and Timon Rabczuk. "Learning and intelligent optimization for material design innovation." In International Conference on Learning and Intelligent Optimization, pp. 358-363. Springer, Cham, 2017.

Torabi, Mehrnoosh, et al. "A hybrid machine learning approach for daily prediction of solar radiation." International Conference on Global Research and Education. Springer, Cham, 2018.

Mosavi, Amirhosein, et al. "Comprehensive review of deep reinforcement learning methods and applications in economics." Mathematics 8.10 (2020): 1640.

Ahmadi, Mohammad Hossein, et al. "Evaluation of electrical efficiency of photovoltaic thermal solar collector." Engineering Applications of Computational Fluid Mechanics 14.1 (2020): 545-565.

Ghalandari, Mohammad, et al. "Flutter speed estimation using presented differential quadrature method formulation." Engineering Applications of Computational Fluid Mechanics 13.1 (2019): 804-810.

Ijadi Maghsoodi, Abteen, et al. "Renewable energy technology selection problem using integrated h-swara-multimoora approach." Sustainability 10.12 (2018): 4481.

Mohammadzadeh S, Danial, et al. "Prediction of compression index of fine-grained soils using a gene expression programming model." Infrastructures 4.2 (2019): 26.

Sadeghzadeh, Milad, et al. "Prediction of thermo-physical properties of TiO₂-Al₂O₃/water nanoparticles by using artificial neural network." *Nanomaterials* 10.4 (2020): 697.

Choubin, Bahram, et al. "Earth fissure hazard prediction using machine learning models." *Environmental research* 179 (2019): 108770.

Emadi, Mostafa, et al. "Predicting and mapping of soil organic carbon using machine learning algorithms in Northern Iran." *Remote Sensing* 12.14 (2020): 2234.

Shamshirband, Shahaboddin, et al. "Developing an ANFIS-PSO model to predict mercury emissions in combustion flue gases." *Mathematics* 7.10 (2019): 965.

Salcedo-Sanz, Sancho, et al. "Machine learning information fusion in Earth observation: A comprehensive review of methods, applications and data sources." *Information Fusion* 63 (2020): 256-272.