Gunning fog-index measurement of customer reviews of the Russian Agricultural Bank

A V Plotnikov$^{1,2}$

$^1$ Perm National Research Polytechnic University, 29 Komsomolsky Av., Perm, 614000, Russia
$^2$ Perm State Agro-Technological University named after Academician D.N. Pryanishnikov, 23, Petropavlovskaia Street, Perm, 614990, Russia

E-mail: plotnikov-av@mail.ru

Abstract. Before purchasing service at a bank, most consumers try to find both negative and positive reviews about this bank on the Internet at the same time. Customer reviews of the Russian Agricultural Bank - an object of research. Gunning fog-index is the primary method used in the study to determine average indices. We used the Textstat package in a Python programming language. We added a correction 0.78 factor applicable to the analysis of the text in Russian. The research results are reviews with a score of 5 points that have a lower fog-index, which allows us to talk about the difference between users.

1. Introduction

IT-world leaders such as Google, Facebook, mail.ru, and other two-side markets are shaping the digital paradigm. The primary efforts of IT companies, software suppliers, and various innovative start-ups are to change the formats of activities through which customers receive the necessary products and services. The business introduces a variety of innovative solutions in the shortest possible time, increasing accessibility, ease of use, and personalizing the proposed list of products based on the individual characteristics of the potential consumer. [1] These shifts are transforming existing business models. [2] Now, companies are guided by digital data and follow the principle of "all goods as a service." Companies are trying to accelerate the digital transformation process to meet growing consumer expectations. [3]

1.1. Digital Transformation in the Banking Sector

Today, promising directions of development of the financial sector of the economy primarily consist of the creation of new banking products created by digitalizing current operating processes, as well as non-banking products created in cooperation with third-party partners. [4; 5] The most urgent task is to transfer the country's financial sector to the next stage of development, joining its structures to Industry 4.0. It will allow to qualitatively increase the level of services rendered, to fully satisfy the modern consumer by innovative technologies that reduce the time for physical maintenance. [6] The last decade was marked by the development of Big Data technology in commercial banks. The structure of the domestic financial market
is undergoing significant changes, the purpose of which is to stimulate industry growth after stagnation against the backdrop of the crisis of 2014-2015, 2020, and leveling today's market volatility. The banking services sector characterizes by a significant increase in competition due to the emergence of new structures (non-banking organizations, Internet banks) that claim to be part of the market and provoke a decrease in the profitability of traditional credit organizations. It forces the latter to continually search for new channels of promotion and sources of income. [7]

1.2. Online reputation management
Today, the role of corporate reputation is an essential element on the path to building a successful business. The development of society, the transition from industrial to informational, the total digitalization of the banking sector has created a situation where the reputation of the bank's brand, the first person of the company, development strategy becomes a fundamental asset of the bank. Researchers and experts are united in their assessments: a good reputation of the company is a necessary factor for investment attractiveness; it increases the prospect of revenue growth. It expands the possibilities of capturing new markets. Reputation has a positive effect on other parameters of the company's business model. [8; 9] In this regard, there is a need to manage the company's reputation, develop strengths, and mitigate the negative aspects of reputation development. Further, in the research part, we will analyze customer reviews to the Agricultural Bank.

2. Research methods
As an object of study, we chose the Russian Agricultural Bank, specializing in financing by agricultural enterprises. We collected 6802 customer reviews about the bank’s service and analyzed the intelligibility of these reviews, estimated from 1 to 5 (where five is the best rating). We will use the Gunning fog-index method (the primary method) to check for “comprehensibility” for the human perception of the reviews left. Gunning fog-index was calculated in a modified Textstat package in a Python programming language. We added a correction factor = 0.78 [10] applicable to the analysis of the text in Russian. Complex words are words that have more than two syllables.

\[ Fog.\ index = 0.4 \cdot \left( \frac{\text{words}}{\text{sentences}} \right) + 100 \cdot \left( \frac{\text{complex words}}{\text{words}} \right) \]

The following fog-index values are accepted:

- 70 and above - no specialized training is required;
- up to 70 - secondary education;
- up to 60 - the intellectual level of training;
- up to 30 - for understanding, you need a scientific degree.

3. Research results
Let's look at the dependence of average indices on estimates (table 1 and figure 1).

| Responses grade | rating | Gunning fog-index |
|-----------------|--------|-------------------|
| 1.0             | 64.915720 |
| 2.0             | 60.958270 |
| 3.0             | 55.683857 |
| 4.0             | 52.982455 |
| 5.0             | 27.099372 |
We observe a slight deviation in the average fog-index for reviews with a rating of four and a significant deviation in reviews with a grade of five. Take a look at the distribution of customer reviews for the Gunning fog-index. Figure number 2 illustrates us a large tail of reviews with fog-index value > 200.

**Figure 1.** The dependence of the average Gunning fog-index on the estimates.

**Figure 2.** Distribution of reviews by fog-index in relative terms.

**Figure 3.** Gunning fog-index clusters distribution.

Clustering Gunning fog-index reviews visualized on figures 4 and 5.

**Figure 4.** Visualization of descriptive statistics on clusters.
We define the procedure for dividing into clusters as follows: `gunning_fog_cluster (x)`: if $x < 30$ return 0 (2595 observations); else if $x < 60$ return 1 (2208 observations); else if $x < 70$ return 2 (433 observations); else return 3 (1566 observations). Let’s see the distribution of reviews by the cluster.

We remove the outliers of the Gunning index values of more than 220 (figure 5). We apply the Seaborn package for improved visualization with the display of observations.

![Visualization of descriptive statistics on clusters (with restrictions) using Seaborn package.](image)

**Figure 5.** Visualization of descriptive statistics on clusters (with restrictions) using Seaborn package.

We construct the correlation matrix according to the available numerical characteristics. We see (figure 6) the following observations: `responses_rating_grade` correlate inversely with `message_length`. If `responses_rating_grade` increases, then `message_length` decreases. It suggests that users try harder when writing negative reviews.

![The correlation matrix of features.](image)

**Figure 6.** The correlation matrix of features.
Figure 7. Visualization of the most dependent variables.

Let's plot in 4 dimensions: responses_message_length, gunning_fog_index, responses_rating_grade (marker color), views (marker size), and look at the observations in another visualization presented.

4. Conclusion
When manipulating user behavior - posting fake reviews, they usually use the least cost: low text quality and fewer letters in messages (customer reviews). Thus, we see interesting observations, reviews with a rating of 5 points have a shorter length and a smaller fog-index. This observation leads to some thoughts. We cannot claim that some reviews may be fake following the law, and we also cannot claim the opposite.

Acknowledgments
The reported study was funded by RFBR, project number 20-310-70042.

References
[1] Alekseeva D A 2019 Digital transformation of the banking system Proc of St Petersburg State University of Economics 119 159-62
[2] Matt C, Hess T and Benlian A 2015 Digital transformation strategies Business and Information Systems Engineering 57 339-43
[3] Akatkin Y M, Karpov O E, Konyavsky V A and Yasinovskaya E D 2017 Digital economy: conceptual architecture of the digital industry ecosystem Business Informatics 42 17-28
[4] Dapp T F 2017 Fintech: the digital transformation in the financial sector Sustainability in a Digital World (Springer Cham) 189-99 doi.org/10.1007/978-3-319-54603-2_16
[5] Cuesta C, Ruesta M, Tuesta D and Uribola P 2015 The digital transformation of the banking industry BBVA Research available at https://www.Bbvaresearch.Com/wp-content/uploads/2015/08/EN_Observatorio_Banca_Digital_vf3.Pdf
[6] Sergeeva A S, Suvorova S D and E A Topina 2019 The development of the banking sector in Russia in the context of the digital economy Int. J. of Humanities and Natural Sciences 12-3 151-5 doi: 1024411/2500-1000-2019-11954

[7] Burkaltseva D D, Kuryanova I V and Tyulin A S 2019 The evolution of banking services in the context of the transformation of socio-economic systems Scientific Herald: finance banks investments 46 63-8

[8] Sparks B A, So K K F and Bradley G L 2016 Responding to negative online reviews: The effects of hotel responses on customer inferences of trust and concern Tourism Management 53 74-85

[9] Cillo V Rialti R Del Giudice M and Usai A 2019 Niche tourism destinations ‘online reputation management and competitiveness in big data era: Evidence from three Italian cases Current Issues in Tourism 1-15

[10] Litvinova T, Zagorovskaya O, Litvinova O and Seredin P 2016 Profiling a set of personality traits of text author: What our words reveal about us. Research in Language 14 0019