Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 190 (2021) 193–203

Procedia Computer Science

www.elsevier.com/locate/procedia

2020 Annual International Conference on Brain-Inspired Cognitive Architectures for Artificial Intelligence: Eleventh Annual Meeting of the BICA Society

The Corruption Perception Index: analysis of dependence on socio-economic indicators

Jenny Domashova*, Anna Politova

National Research Nuclear University "MEPhI", 31 Kashirskoe shosse, Moscow 115409, Russia

Abstract

The article presents the results of applying data analysis methods, in particular cluster analysis and machine learning methods for corruption analysis. The analysis of dependence of the corruption perception index on social and economic indicators in different countries of the world was carried out. Data was collected for the formation of the feature space, the most significant features that have the greatest impact on the corruption perception index were selected, and the countries of the world were clustered according to the selected features. A scheme for conducting a comprehensive analysis to identify the most significant signs and causes affecting the corruption perception index has been developed. Based on the results of clustering, a classification model that can predict the level of corruption in the country based on the values of selected attributes was trained. It is based on the composition of Bagging algorithms with decision trees as basic classifiers. The results of the study can be used by public authorities to develop, organize and adopt appropriate counteraction measures. A software tool based on in the Python language, which allows to perform appropriate analysis using updated data was developed .

Peer-review under responsibility of the scientific committee of the 2020 Annual International Conference on Brain-Inspired Cognitive Architectures for Artificial Intelligence: Eleventh Annual Meeting of the BICA Society

Keywords: corruption; corruption perception index; feature selection; clustering; classification; social indicators; economic indicators

* Corresponding author: Jenny Domashova.
E-mail address: janedom@mail.ru

1877-0509 © 2021 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the 2020 Annual International Conference on Brain-Inspired Cognitive Architectures for Artificial Intelligence: Eleventh Annual Meeting of the BICA Society
10.1016/j.procs.2021.06.024

1. Introduction

Corruption is one of the most significant threats to the stability of any state and its financial situation. Analysis of the sources of corruption, as well as forecasting its volume, is necessary for taking appropriate response and counteraction measures. Corruption is a complex social phenomenon, and the motives of a corrupt behavior are multifaceted and result from interaction at the micro, meso, and macro levels [1]. Due to the increased quality and availability of data, empirical research on corruption has been launched since the late 1990s in order to develop more targeted and more effective anti-corruption policies [2].

The consequences of corruption are manifold.

Corruption increases bureaucratic inefficiency. There is no incentive to rationalize an inefficient system for those who benefit from it by engaging in corrupt activities. Thus, corruption and bureaucratic inefficiency can be a cycle. Firms that pay bribes were found to be more likely to spend more managerial time with bureaucrats. Also the presence of corrupt officials can lead to bureaucratic delays in issuing licenses. [3].

A high level of corruption leads to retardation of the country's growth by affecting the investment climate or the quality of investment. This may happen due to inefficient public investment — even if investment volumes increase in absolute terms, absolute productivity may decrease due to inefficient allocation of funds. Corruption can also lead to a decline in the quality of infrastructure, and this will worsen the country's investment climate. [3].

Countries with higher levels of corruption have lower levels of political and civil liberties. The country's population can also determine political effectiveness through the absence of corruption, non-discrimination, and quality of governance. Thus, corruption undermines confidence in the authorities.

Theoretical arguments have been made for the impact of corruption on economic growth through lower investment levels, lower investment quality, higher levels of indirect taxation, and mismanagement of resources due to distorted incentives.

Corruption increases income inequality and poverty by reducing growth, having a biased tax system, poor - quality social programs, inequality in education, and biased ownership of assets.

Like participation in foreign direct investment, international trade most often requires any form of government license or permit. In countries with a higher level of corruption, the costs associated with obtaining the necessary licenses and permits can be particularly high due to the need to pay bribes, etc. Thus, it is assumed that a higher level of corruption negatively affects the level of international trade.

Corruption may force entrepreneurs to move their funds in the shadow economy. In the article [4] a positive relationship between them was found using data from a number of corruption and shadow economy indices from 145 countries.

Corruption and legalization (laundering) of crime proceeds are closely linked. Corruption is a predicate crime for money laundering, as well as an element of money laundering schemes.

The relevance of this work lies in the formation of an understanding of the origin and sources of such a phenomenon as corruption, as well as in identifying the most significant causes that affect the Corruption Perception Index. This will make it more likely to identify the areas of the economy and social sphere that are most susceptible to corruption. It is also important to predict the corruption risk in the country, which will allow to assess which measures are needed to combat corruption.

2. The Corruption Perception Index

The corruption perception index (CPI) is a corruption index published by the Transparency International. The index has been compiled since 1995 and is published in the form of a rating that compares perceived levels of public sector corruption in 180 countries.

Transparency International is a non-governmental organization dedicated to fighting corruption around the world. The organization defines corruption as the abuse of entrusted power for personal gain, which ultimately harms everyone who depends on the integrity of people in a senior position [5]. Transparency International develops anti -corruption tools and works with other civil society organizations, companies and governments to implement them. Since 1995 Transparency International has published the Corruption Perception Index. The CPI is their most important publication which ranks countries based on their perception of corruption among government officials and politicians. It is a composite index based on various surveys and studies conducted by more than ten independent institutions. Annual interviews with business people and analysts are conducted in various countries, as well as surveys with experts which are held domestically and abroad. The index varies from 0 to 100 (maximum 10 points before 2011), where 100 indicates the lowest level of corruption perception and, therefore, the best possible result. The index is intended to represent alleged corruption. It reflects respondents' views on alleged corruption. This opinion may be based on personal experience of dealing with corruption.

Thus, the Corruption Perception Index uses indices and ratings based on expert surveys but does not provide an understanding of what exactly is the source of corruption. In this article, we study the causes of corruption occurrence, as well as the formation of a list of indicators that can be used to determine the level of corruption in the country.

3. Possible causes of corruption

There are several basic theories which suggest that a higher level of inefficiency in the bureaucratic system increases the level of corruption. The more interactions exist in the regulatory acts, the more frequent they are with officials in the private sector and, consequently, the higher the likelihood that an official may be involved in corrupt activities. In an inefficient bureaucracy, the rules of interaction tend to be less transparent. Both factors imply a higher level of corruption. Excessive government intervention in the economy, namely in the field of regulation, can also contribute to corruption.

A high level of economic freedom, or the freedom to choose how to produce, sell, and use own resources, is associated with a lower level of corruption. With less economic control, it is less likely that corrupt behavior will be considered necessary for doing business.

Press freedom plays an important role in the dissemination of anti-corruption standards, as well as in increasing the potential social costs of public condemnation for corrupt behavior. In addition, civic participation in the form of democracy can combat corruption, as regular elections give the public the opportunity to eliminate corrupt politicians. A research [5] using panel data covering 126 countries from 1980 to 2007 concludes that both democratization and media freedom have a negative impact on corruption.

Corruption increases along with the level of poverty. First, poorer countries are less likely to devote the necessary resources to building an effective legal system. Second, the main motivation for paying bribes in this case would be to gain access to basic public services (such as education, permits, and licenses) for which the state has a monopoly, that is a strong motivation to break the law.

Increasing the level of openness and trade or integration into the global economy should reduce the level of corruption. This is because greater integration can change both the political and economic structure of a country and cultural norms. In addition, increasing the level of free trade will deprive bureaucratic monopolies of some administrative licenses and permits, reducing the likelihood of corrupt behavior.

Increased transparency is associated with a lower level of corruption. The likelihood of detection of offenses increases, as does the responsibility of each decision-maker. Transparency matters only when it is accompanied by free and fair elections, as well as the ability to impose sanctions on corrupt individuals. [4]

There is also a correlation between corruption and salary levels: government officials with higher salaries are less likely to be corrupt. Higher wages reduce involvement in corruption out of necessity (family maintenance), but are unlikely to completely eradicate it, as people can still take bribes out of greed.

4. Analysis of Corruption Perception Index dependance on indicators

The scheme of complex analysis consists of the following sequence of procedures:

- Data collection and generation of a set of features
- Data pre-processing
- Selection of the most informative features
- Clustering of the world's countries based on socio-economic indicators
- Training a model to classify the countries by the level of perceived corruption

4.1. Data collection

The data on the corruption perception index was obtained from the Transparency International website. It is available to download data for the period from 2012 to 2019 for 180 countries in xlsx format (Figure 1).

Corruption Perceptions Index 2019: Score timeseries since 2012

Country	ISO3	Region	CPI score 2019	Rank 2019	Sources 2019	Standard error 2019	CPI score 2018	Rank 2018	Sources 2018	Standard error 2018	CPI score 2017	Rank 2017	Sources 2017	Standard error 2017
New Zealand	NZL	AP	**87**	1	8	2,29	**87**	2	8	2,44	**89**	1	8	2,4
Denmark	DNK	WE/EU	**87**	1	8	2,54	**88**	1	8	2,63	**88**	2	8	2,75
Finland	FIN	WE/EU	**86**	3	8	2,92	**85**	3	8	2,74	**85**	3	8	2,84
Switzerland	CHE	WE/EU	**85**	4	7	1,58	**85**	3	7	1,57	**85**	3	7	1,71
Singapore	SGP	AP	**85**	4	9	2,05	**85**	3	9	1,98	**84**	6	9	2,26
Sweden	SWE	WE/EU	**85**	4	8	1,98	**85**	3	8	2,02	**84**	6	8	2,27
Norway	NOR	WE/EU	**84**	7	7	1,65	**84**	7	8	2,14	**85**	3	8	1,83
Netherlands	NLD	WE/EU	**82**	8	8	2,25	**82**	8	8	2,3	**82**	8	8	2,23
Luxembourg	LUX	WE/EU	**80**	9	7	1,95	**81**	9	7	2,32	**82**	8	6	2,08
Germany	DEU	WE/EU	**80**	9	8	3,31	**80**	11	8	2,49	**81**	12	8	1,87
Iceland	ISL	WE/EU	**78**	11	7	4,63	**76**	14	7	4,33	**77**	13	7	4,38
Canada	CAN	AME	**77**	12	8	2,80	**81**	9	8	2,16	**82**	8	8	1,49
United Kingdom	GBR	WE/EU	**77**	12	8	3,34	**80**	11	8	2,03	**82**	8	8	1,7
Australia	AUS	AP	**77**	12	9	1,32	**77**	13	9	1,27	**77**	13	9	1,4
Austria	AUT	WE/EU	**77**	12	8	1,57	**76**	14	8	1,37	**75**	16	8	1,17
Hong Kong	HKG	AP	**76**	16	8	3,15	**76**	14	8	2,47	**77**	13	7	2,37
Belgium	BEL	WE/EU	**75**	17	7	1,09	**75**	17	8	1,09	**75**	16	8	1,63
Ireland	IRL	WE/EU	**74**	18	7	3,61	**73**	18	7	3,25	**74**	19	7	3,68
Estonia	EST	WE/EU	**74**	18	10	1,21	**73**	18	10	1,44	**71**	21	10	2,21
Japan	JPN	AP	**73**	20	9	3,51	**73**	18	9	2,85	**73**	20	9	2,66
United Arab Emirates	ARE	MENA	**71**	21	8	5,13	**70**	23	8	4,56	**71**	21	7	6,26
Uruguay	URY	AME	**71**	21	7	2,47	**70**	23	7	2,73	**70**	23	7	2,67

Fig. 1. Fragment of the Excel file with the CPI data set

At the next stage, a search was performed on open data sources in order to create a set of features. A data source is considered suitable, if:

1) the data has open access;
2) the indicators are not indices, they are not derived from any indicators or the results of an expert survey;
3) there is information for at least 100 countries;
4) there is data for the required period (2012-2019);
5) priority is given to sources whether data can be downloaded in xlsx or csv formats.

Out of 18 sites, three sources meet the above criteria:

1) Data from the International Monetary Fund (IMF);
2) UN indicators for the Sustainable Development Goals (SDGs of the UN);
3) Data from the Doing Business project of the World Bank.

Then the dataset was formed by adding feature data to CPI data. The final set contains data for 180 countries around the world for the period from 2012 to 2019. Further analysis was carried out using the data of 2018, as this dataset has more features than in 2019, and it has fewer empty values. In total, the set includes 69 features, which can be divided into three main categories (Figure 2):

- economic indicators (18 features),
- social indicators (15 features),
- indicators that assess environments for doing business (36 features).

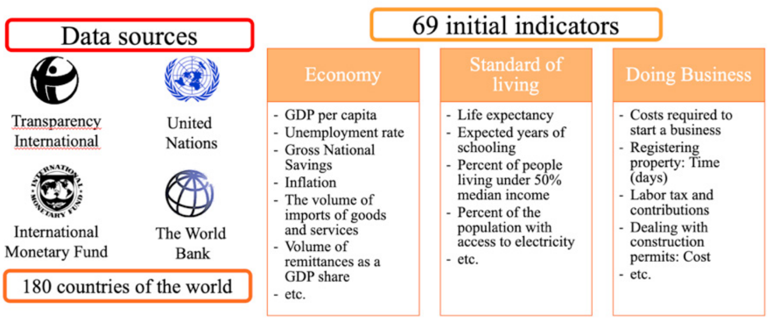


Fig. 2. Creating a list of attributes

4.2. Data pre-processing

The data pre-processing stage includes analyzing and filling in missing values and identifying abnormal values. The following string values indicating the absence of data were identified:

- "No Practice";
- "No corporate income tax";
- "--".

These values have been replaced with the empty ones. Next there was the removal of columns and rows containing more than 80% of empty values. The remaining values were filled with the average values of the characteristic. The obtained data was standardized.

In order to exclude strongly correlated features using the pandas library method DataFrame.corr() correlation matrix of the original features was obtained (Figure 3). By default the method uses the Pearson correlation coefficient, the method returns a dataframe-matrix of pairwise correlations. In Figure 3, the correlation matrix was plotted as a heatmap using the seaborn library: the darker the heatmap cell is, the stronger the pairwise correlation.

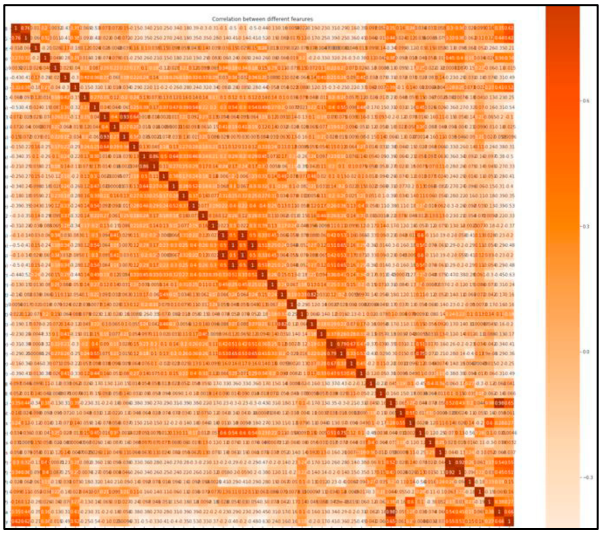


Fig. 3. Heatmap of the original features' correlation

The presence of a close relationship between the features means that one feature from a given pair can be removed without the loss of information. After removal 46 features were left.

4.3. Feature selection

The problem of feature selection consists of the selection of the most significant socio-economic features for reducing training time, simplifying models and interpreting the results.

The feature selection was carried out by three methods created using Python. For classification in the greedy features addition-removal Add and Add-Del methods linear regression model (Linear), support vector machine (SVM), regression based on the k-nearest neighbors algorithm (KNN), regression based on decision trees (Decision Tree) were used. In the feature selection algorithm Breadth-first search (BFS) regression, based on the k-nearest neighbors algorithm, was used for classification.

For the selection of the best model on the test data, the following quality metrics were calculated:

- Mean Squared Error (MSE),
- Root Mean Squared Error (RootMSE),

- Mean Absolute Error (MAE),
- Median Absolute Error (MdAE),
- Coefficient of determination R^2.

Final list included 12 features from all three groups. Table 1 presents the results of the mentioned algorithms and selected features.

Table 1. Results of the features' selection algorithms.

Method	Regression model	Selected features	MSE	$\sqrt{MSE}$	MAE	MdAE	R^2
Add	KNN	Life expectancy, Current account balance (Percent of GDP)	0,028	0,167	0,142	0,122	0,603
Add-Del	KNN	Inflation, end of period consumer prices (Percent change), PT: Other taxes (% of profit), Gross domestic product per capita, current prices (Us dollars), TaB: Time to import: Documentary compliance (hours), PT: Time (hours per year), SB: Cost - Men (% of income per capita), PT: Total tax and contribution rate (% of profit), General government primary net lending/borrowing (Percent of GDP)	0,014	0,12	0,097	0,091	0,797
BFS	KNN	Gross domestic product per capita, current prices (Us dollars), Life expectancy	0,026	0,16	0,121	0,096	0,638

A feature was considered significant if it was selected by different models at least twice. As a result, 12 features were selected from all three groups, the selected features are presented in Table 2.

Table 2. List of selected informative features.

Method	Selected features
Economics	Unemployment rate Inflation General government primary net loans/borrowing Export volume of goods and services
Standard of living	Life expectancy
Business management	Annual growth rate of output per worker Time for dealing with construction permit Time for paperwork for import Cost of paperwork for import Paying taxes: Time Starting a business: Cost Registering property: Time

4.4. Clustering

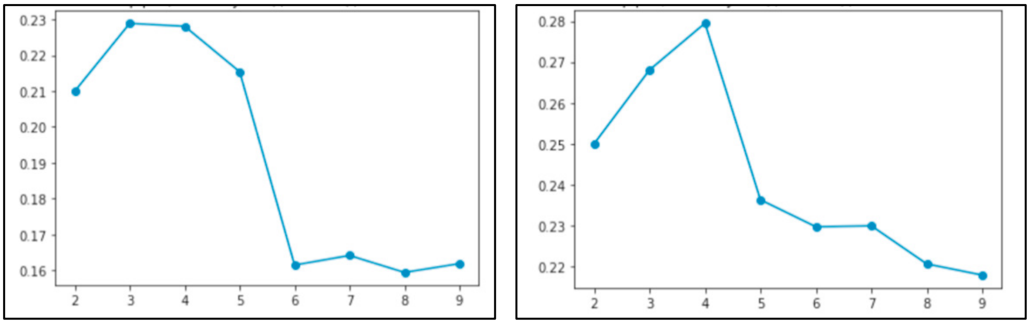
Clustering task is to divide the world countries into homogeneous groups (clusters) according to socio-economic characteristics.

In order to compare results clustering was carried out using two data sets: including selected features and including all features. Three methods were used: hierarchical methods (Complete-linkage clustering and Ward method) and iterative method (K-means).

The Silhouette method is used to select the number of clusters and validate the consistency within clusters. It reflects how much the average distance from one object to other objects in its cluster differs from the average distance to objects in other clusters. The calculation of this coefficient is performed using the silhouette_score ()

method of the sklearn library. The Euclidean distance (metric = euclidean) was chosen as the distance metric between objects.

The values of the Silhouette coefficients on selected features for each of the clustering methods were greater than for clustering on all features. The best value was obtained when clustering on selected features using the K-means clustering (3 clusters) and by the complete-linkage clustering (4 clusters) (figures 4a and 4b).



a) b)

Fig. 4. (a) a plot of the silhouette coefficient values for the K-means clustering; (b) a plot of the silhouette coefficient values for the complete-linkage clustering.

The number of objects and the corruption perception index average for each cluster are presented in Table 3.

Table 3. World countries clustering.

	Complete-linkage clustering		K-means	
	Number of objects	CPI average	Number of objects	CPI average
Cluster 0	121	49	90	51
Cluster 1	6	24	39	19
Cluster 2	41	23	40	41
Cluster 3	1	7	-	-

The K-means clustering identified three clusters characterized by different levels of CPI. The largest cluster has high average CPI value (above 50). The smaller clusters have low and medium CPI. As a result of complete-linkage clustering, the cluster 4 consists of one object. Also, this object is included in a separate cluster while clustering into 3 clusters. Thus, the division into 3 clusters using K-means was used to characterize clusters and for further classification. Clusters' objects are presented in Table 4.

Table 4. K-means clustering on selected features – objects

Cluster	Number of objects	Country
0	90	Oman, Grenada, Italy, Mauritius, Slovakia, Saudi Arabia, Croatia, Rwanda, Costa Rica, Malaysia, South Korea, Cape Verde, Vanuatu, Hungary, Latvia, Georgia, Lithuania, Czech Republic, Belarus, Solomon Islands, Poland, Slovenia, Morocco, Israel, Bulgaria, Qatar, Turkey, Kuwait, Trinidad and Tobago, India, Taiwan, Brunei, Darussalam, Portugal, Serbia, China, Sri Lanka, Seychelles, Panama, Mongolia, Chile, Colombia, Bahrain, Thailand , Philippines, Bhutan, Barbados, Armenia, El Salvador, Peru, Ecuador, United Arab Emirates, Vietnam, Moldova, Pakistan, United States of America, Ukraine, France, Nepal, Kazakhstan, Japan, Estonia, Ireland, Dominican Republic, Bolivia, Laos, Kyrgyzstan, Paraguay, Honduras, Lebanon, Mexico, Hong Kong, Austria, Iceland, Guatemala, Australia, Tajikistan, Nicaragua, Great Britain, Germany, Uzbekistan, Canada, Netherlands, Cambodia, Norway, Sweden, Singapore, Switzerland, Finland, New Zealand, Denmark

1	39	Senegal, Ghana, Burkina Faso, Benin, Gambia, Tanzania, Cote d'Ivoire, Egypt, Ethiopia, Niger, Liberia, Malawi, Mali, Sierra Leone, Togo, Myanmar, Guinea, Papua New Guinea, Comoros, Nigeria, Mauritania, Kenya, Bangladesh, Uganda, Central African Republic, Madagascar, Cameroon, Mozambique, Zimbabwe, Haiti, Chad, Congo, Angola, Venezuela, Burundi, Afghanistan, Guinea-Bissau, Equatorial Guinea, South Sudan
2	40	Namibia, Malta, Russia, Jordan, Saint Lucia, Dominica, Sao Tome and Principe, Spain, Saint Vincent and the Grenadines, Montenegro, Greece, Cyprus, Jamaica, Tunisia, Suriname, South Africa, Botswana, Lesotho, Argentina, Bahamas, Eswatini, Indonesia, Bosnia and Herzegovina, Guyana, North Macedonia, Albania, Brazil, Algeria, Zambia, Uruguay, Gabon, Djibouti, Maldives, Belgium, Iran, Azerbaijan, Luxembourg, Iraq, Sudan, Yemen

Table 5 shows a comparison of the K-Means clustering results on all features and on the selected ones. It can be concluded that the average values of the CPI do not differ much for these clusters. objects. In cluster 0 there are less objects while clustering on all features than while clustering on the selected ones. Cluster 2 includes these objects.

Table 5. Comparison of the results of the K-Means clustering on all the features and on the selected ones.

	Average CPI		Number of objects	
	Selected features	All features	Selected features	All features
Cluster 0	51,5	59,7	91	50
Cluster 1	19,4	21,3	39	51
Cluster 2	41,1	43,5	40	68

The Silhouette coefficient for clustering on the selected features is 0.23, and for clustering on all features - 0.11. It means that clustering on the selected features forms more homogeneous clusters.

Cluster 0 includes the USA, Australia, Canada, European countries and some other countries (Figure 5). These countries have average of living standard and doing business indicators higher than other countries. The CPI average is 53. This cluster has the following average values of indicators:

- low unemployment rates, as well as the costs of starting a business, import and registering property;
- processes of paying taxes and import are not so time-consuming;
- high life expectancy and annual growth rate of output per worker.

Fig. 5. Countries of cluster 0

Cluster 1 includes countries of Africa, South America and Asia (Figure 6). These countries have low levels of economic development and living standards. The CPI average is 19. This cluster has the following average values:

- low life expectancy and annual growth rate of output per worker;
- high unemployment rates, the costs of starting a business, import and registering property;

- processes of paying taxes and import are time-consuming.

Fig. 6. Countries of cluster 1

Cluster 2 includes Russia, Brazil, Argentina and African countries (Figure 7). This cluster is characterized by high property registration time and unemployment rate. The remaining 10 features have average values. The CPI average is 43.

Fig. 7. Countries of cluster 2

4.5. Classification

Then in order to predict the level of corruption in countries classification model was trained. For this purpose, a training sample was formed: it includes the objects that were closest to the centers of each cluster. The size of the training sample was calculated as 20% of the total number of objects and is equal to 33. Elements of the training sample are presented in Table 6.

Table 6. Training sample for classification.

Class	Country
0	Oman, Grenada, Italy, Mauritius, Slovakia, Saudi Arabia, Croatia, Rwanda, Costa Rica, Malaysia, South Korea
1	Senegal, Ghana, Burkina, Faso, Benin, Gambia, Tanzania, Ivory Coast, Egypt, Ethiopia, Niger, Liberia
2	Namibia, Malta, Russia, Jordan, Saint Lucia, Dominica, Sao Tome and Principe, Spain, Saint Vincent and the Grenadines, Montenegro, Greece, Cyprus

Training was carried out on decision trees using 4 ensemble algorithms: Random Forest, AdaBoost, Gradient Boosting, Bagging.

Objects that were not included in the training sample formed the test sample. Best classification quality metrics were obtained when using Bagging. The average value of the accuracy for this algorithm is 0.89. The values of the metrics for each of the classes are shown in Table 7.

Table 7. Classification quality metrics for the Bagging algorithm.

Class	Precision	Recall	f-score	Support
1	0,96	0,91	0,94	80
2	0,76	0,93	0,84	28
3	0,85	0,79	0,81	28

Using sklearn.metrics.confusion_matrix was obtained the number of test sample objects which were assigned by the classification algorithm to each of the classes (Table 8). The number of incorrectly classified objects is insignificant

Table 8. Bagging. Confusion matrix.

		Class formed by the algorithm		
		1	2	3
Original class number	1	73	4	3
	2	1	26	1
	3	2	4	22

During classification with this algorithm sklearn.ensemble library's function BaggingClassifier was given the following parameters: base_estimator=None, bootstrap=True, bootstrap_features=False, max_features=1.0, max_samples=1.0, n_estimators=10, n_jobs=None, oob_score=False, random_state=0, verbose=0, warm_start=False.

5. Conclusion

Corruption Perception Index formed by the organization Transparency International assesses and ranks the countries, based on experts' and business leaders' opinions on prevalence of corruption in the country. It is a composite index based on 13 surveys and corruption assessments compiled by various reputable organizations. Within the framework of the research, a scheme for studying dependance of corruption on social and economic indicators in different countries is proposed. A list of data sources is proposed, a set of informative indicators that have greater impact on the level of corruption is built, clustering of the world countries on the selected features is performed.

Characteristics of the cluster with the highest average level of corruption perception can be used to identify areas that directly reduce corruption. In particular, an increase in the values of indicators such as life expectancy and the annual growth rate of output per worker, as well as a decrease in the unemployment rate, the cost of launching a business, registering property and imports, in the time spent on paying taxes and processing documents on import can lead to the reduction of corruption in countries.

The classification model was trained which can later be used to assess the risk of corruption in the country. It can help to choose how extensive the provided measures should be in order to combat corruption. This information will be useful for national risk assessment of money laundering and terrorist financing.

The results of this work can be used to develop the countries' legislation in the field of combating corruption, identifying problems in bureaucratic systems and in developing anti-corruption policies, as well as policies for economic and social development.

Acknowledgements

This work is supported by the National Research Nuclear University "MEPhI".

References

[1] Bicchieri, Ganegonda (2017) "Determinants of corruption: a socio-psychological analysis." in *Thinking About Bribery, Neuroscience, Moral Cognitionand the Psychology of Bribery* Nichols, Robertson.

[2] Lambsdorff, Schulze (2015) "What can we know about corruption? A very short history of corruption research and a list of what we should aim for. "*Journal of Economics and Statistics* **235 (2)**: 100-114.

[3] Dimant, Tosato (2018) "Causes and effects of corruption: What has past decade's empirical research taught us? A survey. "*Journal of economic surveys* **32(2):** 335-356.

[4] Dell'Anno, R., Teobaldelli, D. (2016) "Keeping both corruption and the shadow economy in check: the role of decentralization "*International Tax and Public Finance,* vol. **22(1)**: 1-40.

[5] Bhattacharyya, Hodler (2015) "Media freedom and democracy in the fight against corruption" *European Journal of Political Economy*, **39 (C):** 13-24.

Authors

1. Jenny Domashova, Ph.D. (Econ.), Associate Professor, Institute of Financial Technology and Economic Security, NRNU "MEPhI", Moscow, Russia, janedom@mail.ru, ORCiD 0000-0003-1987-8553
2. Anna Politova, student, Institute of Financial Technology and Economic Security, NRNU "MEPhI", Moscow, Russia, politova-av@yandex.ru, ORCiD 0000-0002-6656-6790.