

Министерство образования и науки Российской Федерации  
Высшая школа управления и инноваций МГУ им. М.В. Ломоносова (ВШУИ МГУ)  
II Международная научно-практическая конференция «Инновационная экономика и менеджмент: методы и технологии»

*Тема доклада: «РАЗРАБОТКА МЕТОДИКИ ПО АВТОМАТИЗАЦИИ ПРОЦЕССА СБОРА, ОБРАБОТКИ И АНАЛИЗА ДАННЫХ О ДЕЯТЕЛЬНОСТИ КРЕДИТНЫХ ОРГАНИЗАЦИЙ В ЦЕЛЯХ ПРОГНОЗИРОВАНИЯ РИСКОВ ЭКОНОМИЧЕСКОЙ БЕЗОПАСНОСТИ»*

---

ПРИКАЗЧИКОВА А. С., АСЛАНОВ Р. Э.

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЯДЕРНЫЙ УНИВЕРСИТЕТ НИЯУ «МИФИ»

26 апреля 2017 г., Москва

# Обзор проблемной области и постановка задачи

---

Проблема оценки благонадежности субъектов экономического сектора является одной из актуальных в сфере финансового мониторинга.

Целью проведенного исследования является апробация одного из методов машинного обучения в целях выявления рисков финансовой устойчивости кредитных организаций.



# Формы финансовой отчетности кредитных организаций



На сайте Центрального банка России представлены следующие формы банковской отчетности:

- Данные оборотной ведомости по счетам бухгалтерского учёта, форма 101
- Отчёт о финансовых результатах, форма 102
- Расчёт собственных средств (капитала) («Базель III») Расчёт собственных средств (капитала)
- Информация об обязательных нормативах, форма 135
- Консолидированная финансовая отчетность
- Финансовая отчетность

Код формы по ОКУД 0409101  
тыс. рублей

Номер счета второго порядка	Входящие остатки			Обороты за отчетный период					
				по дебету			по кредиту		
	в рублях	ин. вал., драг. металлы	итого	в рублях	ин. вал., драг. металлы	итого	в рублях	ин. вал., драг. металлы	итого
1	2	3	4	5	6	7	8	9	
<b>А. Балансовые счета</b>									
<b>Актив</b>									
20202	665 705	1 010 110	1 675 815	11 198 491	3 926 370	15 124 861	11 053 602	4 078 103	
20208	65 597	0	65 597	259 398	0	259 398	254 995	0	
20209	10 152	0	10 152	6 956 172	1 531 901	8 488 073	6 935 173	1 531 901	
20302	0	37 139	37 139	0	5 975	5 975	0	3 570	
20303	0	133	133	0	45	45	0	17	
20308	0	1 447	1 447	0	330	330	0	249	
20311	0	4 083	4 083	0	646	646	0	560	
20319	0	14	14	0	24	24	0	23	
30102	752 074	0	752 074	38 940 972	0	38 940 972	38 574 738	0	
30110	60 927	91 003	151 930	372 244	870 087	1 242 331	396 560	871 331	
30114	0	360 061	360 061	0	1 698 059	1 698 059	0	1 393 655	
30202	274 365	0	274 365	14 925	0	14 925	0	0	
30204	128 739	0	128 739	4 883	0	4 883	0	0	
30210	0	0	0	48 864	0	48 864	48 864	0	

Рис. № 1. Форма банковской отчетности

# Формирование матрицы исходных данных

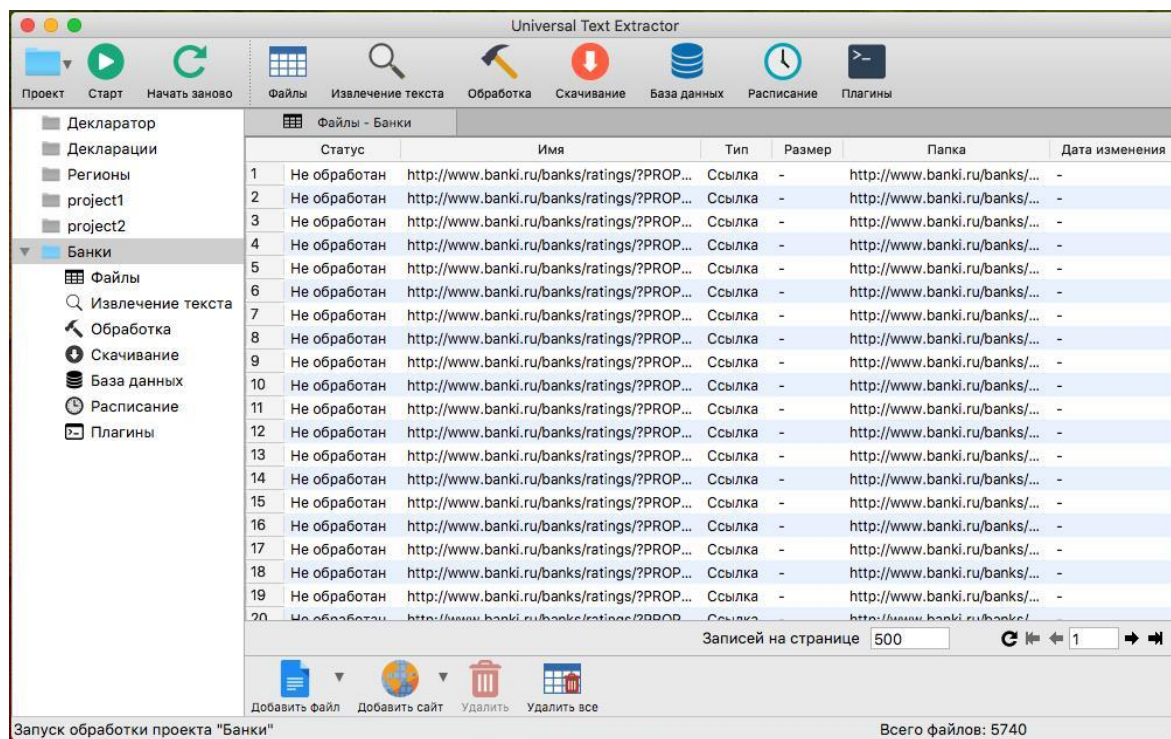


Рис. № 2. Парсер данных о деятельности КО

```
1 # -*- coding: utf-8 -*-
2 import csv
3 import pandas
4
5
6 # Load indicators
7 indicators = {}
8 with open('indicators.csv', 'r', encoding='utf-8') as input_file:
9     for row in csv.reader(input_file, delimiter=','):
10         key = int(row[1])
11         indicators[key] = indicators[int(row[2])] + ' -> ' + row[0] if row[2] else row[0]
12
13 # Load data
14 df = pandas.read_csv('data.csv', sep=';', encoding='utf-8')
15 # Delete unnecessary symbols
16 df['Банк'] = df['Банк'].replace(to_replace='\x14 ', value='', regex=True)
17 df['Значение'] = df['Значение'].replace(to_replace='\x12', value='', regex=True) \
18                                     .replace(to_replace=',', value='.', regex=True) \
19                                     .astype(float)
20 # Create pivot table
21 df = pandas.pivot_table(df, index=['Дата', 'Банк', 'Лицензия', 'Регион'],
22                         columns='Показатель',
23                         values='Значение')
24 # Change column names and order
25 df.rename(columns={k: indicators[k] for k in df.columns}, inplace=True)
26 df.sort_index(axis=1, inplace=True)
27 df.reset_index(inplace=True)
28 # Save result to new files
29 for date in df['Дата'].unique():
30     df[df['Дата'] == date].to_csv(date + '.csv', encoding='utf-8-sig', sep=';', index=False)
31
```

Рис. № 3. Код на языке Python. Подготовка информации к анализу

# Результат формирования рабочей таблицы с данными о деятельности кредитных организаций

The screenshot displays the STATISTICA 64 interface with a spreadsheet titled 'Factor scores for 5'. The spreadsheet contains 31 rows of data, each representing a credit organization. The columns are organized as follows: Column 1 is 'Наименование КО' (Organization Name); Column 2 is 'Лицензия' (License); Column 3 is 'Отзыв' (Review); Column 4 is 'Регион' (Region); Columns 5-13 are labeled 'FACTOR1' through 'FACTOR8' and 'FACTOR10' through 'FACTOR12', representing the factor scores for five different factors. The data is as follows:

	1	2	3	4	5	6	7	8	9	10	11	12	13
	Наименование КО	Лицензия	Отзыв	Регион	FACTOR1	FACTOR2	FACTOR3	FACTOR4	FACTOR5	FACTOR6	FACTOR7	FACTOR8	FACTOR9
1	Викинг	2	0	Санкт-Петербург и обл.	-0.04595	-0.02622	-0.07261	-0.06115	-0.14258				
2	Кредит-Москва	5	1	Москва и обл.	-0.03943	-0.02472	-0.06996	-0.06208	-0.14460				
3	Примтеркомбанк	21	0	Приморский край	-0.04745	-0.02937	-0.07280	-0.06126	-0.14402				
4	Александровский	53	0	Санкт-Петербург и обл.	-0.04210	-0.02630	-0.06625	-0.05812	-0.14537				
5	Темпбанк	55	0	Москва и обл.	-0.03132	-0.04281	-0.06538	-0.06034	-0.11299				
6	Кольцо Урала	65	0	Свердловская обл.	-0.02103	-0.01510	-0.06795	-0.06888	-0.14985				
7	Саммит Банк	85	0	Приморский край	-0.04550	-0.03230	-0.07292	-0.05965	-0.14334				
8	Кемсоцинбанк	96	0	Кемеровская обл.	-0.04518	-0.02981	-0.07217	-0.06075	-0.14372				
9	БКС — Инвестиционный Банк	101	0	Новосибирская обл.	-0.07789	-0.05577	0.15797	-0.09631	-0.17604				
10	Расчетно-Кредитный Банк	103	0	Москва и обл.	-0.03356	-0.06842	-0.05557	-0.07636	-0.00482				
11	Сельмашбанк	106	0	Ростовская обл.	-0.04644	-0.03237	-0.07260	-0.06055	-0.14349				
12	Континент Финанс	149	1	Москва и обл.	-0.04677	-0.03147	-0.07250	-0.06051	-0.14295				
13	РН Банк	170	0	Москва и обл.	-0.00385	-0.00918	-0.09316	-0.00829	-0.17592				
14	Международный Банк Санкт-Петербурга (МБСП)	197	0	Санкт-Петербург и обл.	-0.05816	0.02314	0.03316	-0.07526	-0.11814				
15	Экспресс-Кредит	210	1	Москва и обл.	-0.03922	-0.02845	-0.07249	-0.06124	-0.14567				
16	Зенит Сочи	232	0	Краснодарский край	-0.04501	-0.02959	-0.07202	-0.06107	-0.14468				
17	Спецстройбанк	236	0	Москва и обл.	-0.04597	-0.03022	-0.07249	-0.06089	-0.14340				
18	Саратов	330	0	Саратовская обл.	-0.04506	-0.03193	-0.07257	-0.06043	-0.14311				
19	Единая Расчетная Палата	384	0	Москва и обл.	-0.04693	-0.03130	-0.07256	-0.06048	-0.14291				
20	Идея Банк	430	1	Краснодарский край	-0.04102	-0.03306	-0.06523	-0.06298	-0.14490				
21	Бест Эффортс Банк	435	0	Москва и обл.	-0.05168	-0.01362	-0.04705	-0.06842	-0.14691				
22	Камский Коммерческий Банк	438	0	Татарстан	-0.03667	-0.03647	-0.07356	-0.05728	-0.13459				
23	Балаково-Банк	444	0	Саратовская обл.	-0.04655	-0.03122	-0.07279	-0.05998	-0.14332				
24	Крыловский	456	1	Краснодарский край	-0.04189	-0.03348	-0.06991	-0.06111	-0.14377				
25	Новопокровский	467	0	Краснодарский край	-0.04214	-0.03556	-0.07703	-0.04678	-0.13768				
26	Русский Национальный Банк	469	0	Ростовская обл.	-0.03855	-0.03500	-0.07469	-0.05149	-0.14489				
27	Енисей	474	1	Красноярский край	-0.04520	-0.02334	0.01927	-0.07908	-0.14357				
28	Кубаньторгбанк	478	0	Краснодарский край	-0.04985	-0.03552	-0.07205	-0.06252	-0.14692				
29	Татсоцбанк	480	0	Татарстан	-0.04790	-0.02538	-0.07417	-0.06411	-0.15393				
30	Ростфинанс	481	0	Ростовская обл.	-0.04290	-0.03028	-0.07298	-0.05556	-0.14221				
31	Донкомбанк	492	0	Ростовская обл.	-0.04270	-0.03325	-0.07403	-0.05375	-0.14523				

Рис. № 4. Матрица исходных данных

# Методы машинного обучения. Метод k-ближайших соседей (k-nearest neighbors)

Для произвольного объекта  $x \in \{x_i\}_{i=1}^m$  расположим объекты обучающей выборки  $x_i$  в порядке возрастания расстояний до  $x$ :

$$\rho(x, x_{1;x}) \leq \rho(x, x_{2;x}) \leq \dots \leq \rho(x, x_{m;x}),$$

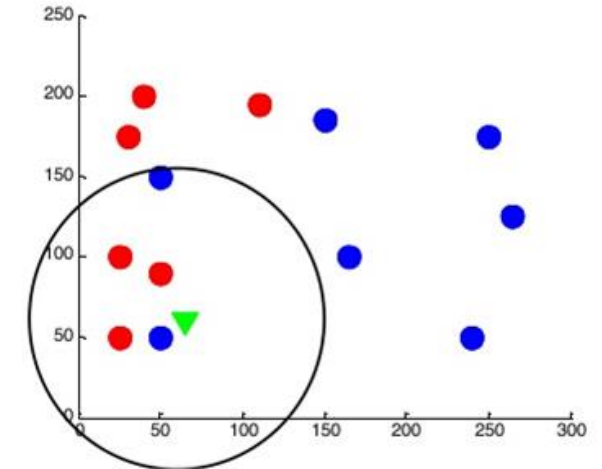
где через  $x_{i;x}$  обозначается тот объект обучающей выборки, который является  $i$ -м соседом объекта  $x$ . Аналогичное обозначение введём и для ответа на  $i$ -м соседе:  $y_{i;x}$ .

Таким образом, произвольный объект  $x$  порождает свою перенумерацию выборки.

В наиболее общем виде алгоритм ближайших соседей:

$$a(x) = \arg \max_{y \in Y} \sum_{i=1}^m [x_{i;x} = y] w(i, x),$$

где  $w(i, u)$  – заданная весовая функция, которая оценивает степень важности  $i$ -го соседа для классификации объекта  $u$ . Естественно полагать, что эта функция неотрицательна и не возрастает по  $i$ .



# Результаты применения метода $k$ -ближайших соседей на полученных данных

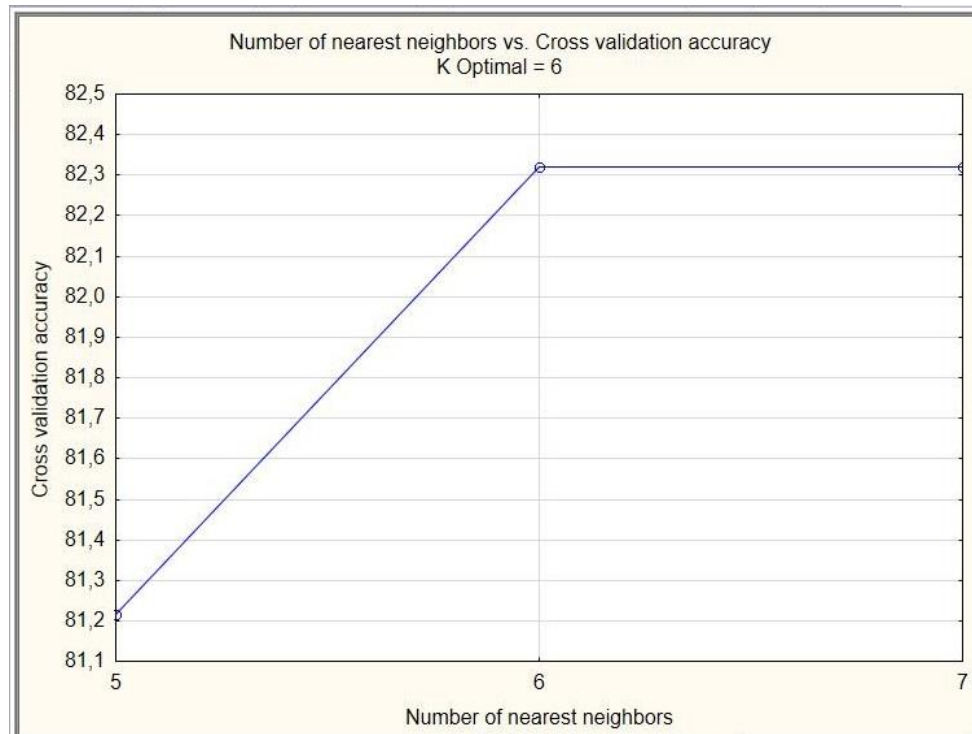


Рис. № 5. Оптимальное число ближайших соседей

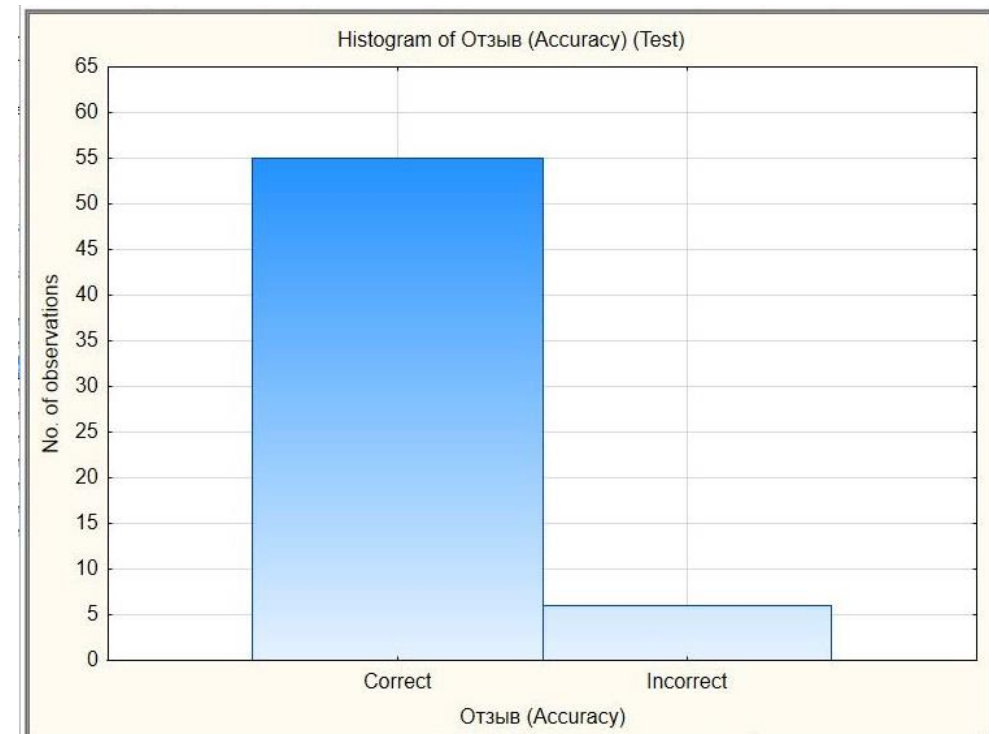


Рис. № 6. Достоверность классификации тестовой выборки

# *Выводы по проделанному анализу*



1. Применение метода машинного обучения k-ближайших соседей к задаче выявления рисков нарушения экономической безопасности кредитных организаций позволяет с высокой степенью вероятности оценить финансовое состояние того или иного банка.
2. Предложенное решение имеет прикладное значение, так как позволяет выявлять потенциально проблемные кредитные организации.
3. Место субъективных экспертных оценок заменяют точные объективные количественные оценки.





***Спасибо за внимание!***