

Эффективность применения метода SVM в задаче определения рентабельных организаций

А. Д. Черемухин¹✉, А. А. Шамин¹, М. О. Колбанев², В. В. Цехановский²

¹ Нижегородский государственный инженерно-экономический университет, Княгинино, Россия

² Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина), Санкт-Петербург, Россия

✉ ngie.u.cheremuhin@yandex.ru

Аннотация. В статье рассматриваются разные способы применения метода опорных векторов при решении задач сельскохозяйственного производства, анализируется эффективность решения задачи поиска факторов, влиявших на прибыльность/убыточность организаций Нижегородской области, занимающихся сельскохозяйственным производством. В число рассмотренных факторов вошли показатели, характеризующие размер организации, объем экономических ресурсов в ней. С помощью процедуры рекурсивного извлечения были отобраны 5 наиболее влиявших на получение прибыли факторов за период 2007–2017 гг. по каждому из 12 рассмотренных видов сельскохозяйственной продукции. Для каждой подвыборки был создан ряд моделей с использованием метода опорных векторов (с разными ядрами и параметрами), выбрана оптимальная и оценена ее точность. В результате был построен рейтинг рассмотренных факторов по важности их влияния на прибыльность/убыточность организации в зависимости от видов производимой продукции.

Ключевые слова: метод опорных векторов, сельскохозяйственные организации, отбор признаков, прибыльность, рекурсивное извлечение, рейтинг, оптимизация параметров, точность моделей

Для цитирования: Эффективность применения метода SVM в задаче определения рентабельных организаций / А. Д. Черемухин, А. А. Шамин, М. О. Колбанев, В. В. Цехановский // Изв. СПбГЭТУ «ЛЭТИ». 2023. Т. 16, № 4. С. 30–45. doi: 10.32603/2071-8985-2023-16-4-30-45.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Original article

The Effectiveness of the SVM Method in the Task of Determining Profitable Organizations

A. D. Cheremuhin¹✉, A. A. Shamin¹, M. O. Kolbanev², V. V. Tsehanovsky²

¹ Nizhny Novgorod Engineering-Economic State University, Knyaginino, Russia

² Saint Petersburg Electrotechnical University, Saint Petersburg, Russia

✉ ngie.u.cheremuhin@yandex.ru

Abstract. The article consider various options of applying the method of support vectors (SVM) for solving problems in the field of agricultural production. In particular, the effectiveness of SVM is analyzed in determining factors affecting the profitability/unprofitability of organizations in the Nizhny Novgorod region engaged in agricultural production. The considered factors included the size of the organization and the amount of available economic resources. Using a recursive extraction procedure, five most decisive factors that affected the profit obtained in 2007–2017 were selected for each of the 12 types of agricultural products considered. For

each subsample, models were built using SVM (with different cores and parameters), the most optimal among which was selected and its accuracy was evaluated. As a result, the considered factors were ranked according to their impact on the profitability/unprofitability of an organization, depending on the types of products produced.

Keywords: support vector machine, agricultural organizations, feature selection, profitability, recursive extraction, rating, parameter optimization, model accuracy

For citation: The Effectiveness of the SVM Method in the Task of Determining Profitable Organizations / A. D. Cheremuhin, A. A. Shamin, M. O. Kolbanev, V. V. Tsehanovsky // LETI Transactions on Electrical Engineering & Computer Science. 2023. Vol. 16, no. 4. P. 30–45. doi: 10.32603/2071-8985-2023-16-4-30-45.

Conflict of interest. The authors declare no conflicts of interest.

Введение. По прогнозам центра инвестирования ARK (Active Research Knowledge) Investment экономическая ценность технологии глубокого изучения составляет около 30 трлн \$, причем современная освоенность ее на сегодняшний момент не превышает 2 трлн \$. Можно с уверенностью говорить, что использование данной технологии позволит создать большое количество практикоориентированных приложений, в том числе и в области сельского хозяйства.

Большой объем технических и аналитических задач, связанных с сельскохозяйственным производством, предполагает решение задачи классификации на уже фиксированные классы (определение по фотографии стадии зрелости продукции растениеводства, степени заболеваемости, предсказание сортности продукции, определение убыточности/прибыльности вероятностного контрагента и т. д.).

При этом очень часто данные задачи решаются с помощью метода опорных векторов и его модификаций).

Постановка задачи. SVM (support vector machine, метод опорных векторов) в классическом его применении – это линейный классификатор, относящийся к методам обучения с учителем, основанный на идее перевода исходных векторов в пространство более высокой размерности и поиска разделяющей гиперплоскости с максимальным зазором в этом пространстве или снижения величины функции ошибки. При этом данный метод может превращаться в нелинейный классификатор при применении разных нелинейных функций перевода в пространство большей размерности.

В настоящее время существует много исследований, посвященных применению метода опорных векторов и разработке различных его модификаций. В недавних работах предлагались новые подходы к разработке ядер (функций перевода в пространства

большой размерности), о чем пишут и зарубежные ученые [1]. Анализ последних публикаций позволяет выделить следующие новые идеи:

- было предложено использование автокорреляционной функции в качестве ядра [2];
- были использованы ортогональные многочлены и объединение классических ядер с функцией Эрмита [3];
- исследовались перспективы и точность гибридных ядер – объединения полиномиального и RBF-ядер [4];
- в [5] исследовалась точность классификации последовательности звуков с использованием экспоненциального ядра Холлингера с использованием вероятностных параметров;
- авторы [6] исследовали построение смешанной ядерной функции для более точной идентификации по классам наземных объектов по данным лидара самолета;
- было создано новое ядро, основанное на применении технологий KSVM в оптимизации совместно с технологией непараметрической поддержки iKN-SVM для реализации методики постепенного обучения и решения проблемы обнаружения лиц [7].

Кроме того, большую популярность приобретает смешение метода опорных векторов и иных векторов. В исследовании [8] была предложена концепция WSVM (Weighted SVM, взвешенный SVM) с использованием вейвлетов для прогнозирования значений энергии ветра. Другим коллективом ученых [9] был предложен многоуровневый алгоритм классификации на несколько классов, основанный на применении SVM, параметры которого определялись генетическими алгоритмами. Предпринимаются также попытки оптимизации метода. Предлагалась методика ускорения процесса обучения на основе технологии фикси-

рованного окна, а в ряде работ исследовалась эффективность применения роевых [10]–[14] или эволюционных [15]–[19] алгоритмов для настройки параметров метода опорных векторов. Еще одним поднаправлением данного типа исследований можно считать разработку дерева решений SVM (Decision Tree SVM; DTSVM) [20], активно продолжают исследования в области разработки и применения нечеткого SVM (Fuzzy SVM; FSVM) [21], [22].

Согласно сообщениям, реализован алгоритм SVM-Boosting на основе марковской передискретизации [23], разработаны различные методики соединения SVM, бустинга и бэггинга [24]–[26].

Анализ точности применения метода LS-SVM (метода опорных векторов с использованием ме-

тода наименьших квадратов; Least Squares – Support Vector Machine) показал большую его применимость относительно нейронных сетей к прогнозированию температуры воздуха [27].

В [28] была использована модифицированная регрессионная модель SVM для моделирования качества воздуха в городских областях. Это исследование может быть перенесено и на сельские территории для предсказания состава атмосферы на полях. Индонезийские ученые применяли регрессию на основе метода опорных векторов для предсказания урожайности риса [29], [30].

В целом метод опорных векторов устойчив и надежен для задач классификации, что и послужило основой его выбора для решения задачи определе-

Табл. 1. Точность моделей определения прибыльности сельскохозяйственных организаций, производящих зерно, построенных на основании метода опорных векторов с разными ядрами
 Tab. 1. Accuracy of models for determining the profitability of agricultural organizations producing grain based on the method of support vectors with different cores

Кол-во влияющих факторов	Применяемое ядро сглаживания	Точность идентификации убыточных организаций	Точность идентификации прибыльных организаций	Общая точность модели	Скорректированная точность модели
1	Линейное	0.000	1.000	0.582	0.284
	Полиномиальное	0.000	1.000	0.582	0.284
	RBF-ядро	0.000	1.000	0.582	0.284
	Сигмоидное	0.001	0.990	0.578	0.284
2	Линейное	0.016	0.996	0.587	0.290
	Полиномиальное	0.302	0.907	0.654	0.424
	RBF-ядро	0.365	0.888	0.669	0.460
	Сигмоидное	0.550	0.675	0.623	0.569
3	Линейное	0.019	0.994	0.588	0.292
	Полиномиальное	0.325	0.911	0.666	0.438
	RBF-ядро	0.307	0.904	0.655	0.427
	Сигмоидное	0.507	0.585	0.552	0.522
4	Линейное	0.079	0.970	0.598	0.316
	Полиномиальное	0.336	0.904	0.667	0.444
	RBF-ядро	0.503	0.844	0.702	0.550
	Сигмоидное	0.508	0.637	0.583	0.532
5	Линейное	0.075	0.972	0.598	0.314
	Полиномиальное	0.285	0.929	0.660	0.416
	RBF-ядро	0.393	0.906	0.692	0.479
	Сигмоидное	0.384	0.687	0.561	0.451

ния факторов, влияющих на рентабельность сельскохозяйственных организаций, и их динамики во времени, а также изменчивости их в зависимости от производимых продуктов.

Результативным показателем в данном исследовании выступал факт наличия прибыльности у сельскохозяйственной организации. Предиктора-ми, по материалам исследований других авторов, выступает количество ресурсов в организации [31], [32]. Нами в качестве ресурсов были взяты количество работников, посевная площадь, условное поголовье, количество, мощность и стоимость сельскохозяйственной техники, а методика соответствует изложенной в [33].

Математическая модель. Все расчеты проводились с использованием языка R и пакета e1071.

Будем рассматривать результаты применения описанного алгоритма к данным о рентабельности сельскохозяйственных организаций по отдельным группам организаций, производящих отдельные виды продукции. Рассмотрим организации, производящие зерно, с пятью наиболее значимыми факторами для данной группы организаций:

- общее количество убранных площадей, га;
- всего пашни, га;
- среднемесячная зарплата, р.;

Табл. 2. Точность моделей определения прибыльности сельскохозяйственных организаций, производящих рапс, построенных на основании метода опорных векторов с разными ядрами
Tab. 2. Accuracy of models for determining the profitability of agricultural organizations producing rapeseed based on the method of support vectors with different cores

Кол-во влияющих факторов	Применяемое ядро сглаживания	Точность идентификации убыточных организаций	Точность идентификации прибыльных организаций	Общая точность модели	Скорректированная точность модели
1	Линейное	0.000	1.000	0.650	0.310
	Полиномиальное	0.010	1.000	0.654	0.315
	RBF-ядро	0.010	1.000	0.654	0.315
	Сигмоидное	0.010	0.982	0.642	0.313
2	Линейное	0.000	1.000	0.650	0.310
	Полиномиальное	0.010	1.000	0.654	0.315
	RBF-ядро	0.000	1.000	0.650	0.310
	Сигмоидное	0.053	0.942	0.631	0.327
3	Линейное	0.000	1.000	0.650	0.310
	Полиномиальное	0.075	0.994	0.672	0.341
	RBF-ядро	0.086	0.994	0.676	0.346
	Сигмоидное	0.000	1.000	0.650	0.310
4	Линейное	0.000	1.000	0.650	0.310
	Полиномиальное	0.075	0.994	0.672	0.341
	RBF-ядро	0.000	1.000	0.650	0.310
	Сигмоидное	0.000	0.988	0.642	0.309
5	Линейное	0.000	1.000	0.650	0.310
	Полиномиальное	0.086	1.000	0.680	0.346
	RBF-ядро	0.000	1.000	0.650	0.310
	Сигмоидное	0.000	0.982	0.639	0.309

Табл. 3. Точность моделей определения прибыльности сельскохозяйственных организаций, производящих картофель, построенных на основании метода опорных векторов с разными ядрами
Tab. 3. Accuracy of models for determining the profitability of agricultural organizations producing potatoes based on the method of support vectors with different cores

Кол-во влияющих факторов	Применяемое ядро сглаживания	Точность идентификации убыточных организаций	Точность идентификации прибыльных организаций	Общая точность модели	Скорректированная точность модели
1	Линейное	0.000	1.000	0.694	0.325
	Полиномиальное	0.044	0.997	0.706	0.342
	RBF-ядро	0.051	0.997	0.708	0.345
	Сигмоидное	0.000	1.000	0.694	0.325
2	Линейное	0.051	0.997	0.708	0.345
	Полиномиальное	0.068	0.997	0.713	0.352
	RBF-ядро	0.238	0.964	0.742	0.428
	Сигмоидное	0.000	1.000	0.694	0.325
3	Линейное	0.051	0.997	0.708	0.345
	Полиномиальное	0.051	0.997	0.708	0.345
	RBF-ядро	0.337	0.952	0.764	0.479
	Сигмоидное	0.000	1.000	0.694	0.325
4	Линейное	0.051	0.997	0.708	0.345
	Полиномиальное	0.068	0.998	0.714	0.352
	RBF-ядро	0.450	0.937	0.788	0.545
	Сигмоидное	0.221	0.812	0.632	0.403
5	Линейное	0.051	0.997	0.708	0.345
	Полиномиальное	0.078	0.997	0.716	0.356
	RBF-ядро	0.535	0.946	0.820	0.601
	Сигмоидное	0.078	0.950	0.684	0.353

- условное поголовье скота, гол.;
- мощность техники, л. с.

Результаты применения метода SVM с различными функциями представлены в табл. 1.

Соответственно, лучшую точность по обоим критериям показывает модель с RBF-ядром и четырьмя значимыми факторами, доля правильно классифицированных объектов – 70.2 %.

Рассмотрим организации, производящие рапс, с пятью наиболее значимыми для них факторами:

- всего пашни, га;
- всего сельскохозяйственных угодий, га;
- среднегодовая численность работников, чел.;
- полная себестоимость продукции животноводства, тыс. р.;
- стоимость основных средств, тыс. р.

Результаты применения метода SVM с различными функциями представлены в табл. 2.

Соответственно, лучшую точность по обоим критериям показывает модель с полиномиальным ядром и пятью значимыми факторами, доля правильно классифицированных объектов – 68.0 %.

Рассмотрим организации, производящие картофель, с пятью наиболее значимыми для них факторами:

- условное поголовье скота, гол.;
- количество сельскохозяйственной техники, шт.;
- стоимость основных средств, тыс. р.;
- среднегодовая численность работников, чел.;
- общее количество убранных площадей, га.

Результаты применения метода SVM с различными функциями представлены в табл. 3.

Соответственно, лучшую точность по обоим критериям показывает модель с RBF-ядром и пятью значимыми факторами, доля правильно классифицированных объектов – 82.0 %.

Рассмотрим организации, производящие сахарную свеклу, с пятью наиболее значимыми для них факторами:

- общее количество убранных площадей, га;
- полная себестоимость продукции животноводства, тыс. р.;
- стоимость основных средств, тыс. р.;
- всего пашни, га;
- всего сельскохозяйственных угодий, га.

Табл. 4. Точность моделей определения прибыльности сельскохозяйственных организаций, производящих сахарную свеклу, построенных на основании метода опорных векторов с разными ядрами

Tab. 4. Accuracy of models for determining the profitability of agricultural organizations producing sugar beet based on the method of support vectors with different cores

Кол-во влияющих факторов	Применяемое ядро сглаживания	Точность идентификации убыточных организаций	Точность идентификации прибыльных организаций	Общая точность модели	Скорректированная точность модели
1	Линейное	0.000	1.000	0.725	0.333
	Полиномиальное	0.000	1.000	0.725	0.333
	RBF-ядро	0.000	1.000	0.725	0.333
	Сигмоидное	0.000	1.000	0.725	0.333
2	Линейное	0.454	0.965	0.825	0.555
	Полиномиальное	0.454	0.965	0.825	0.555
	RBF-ядро	0.454	0.965	0.825	0.555
	Сигмоидное	0.000	1.000	0.725	0.333
3	Линейное	0.454	0.965	0.825	0.555
	Полиномиальное	0.454	1.000	0.850	0.556
	RBF-ядро	0.454	0.965	0.825	0.555
	Сигмоидное	0.363	0.965	0.800	0.501
4	Линейное	0.454	0.965	0.825	0.555
	Полиномиальное	0.454	1.000	0.850	0.556
	RBF-ядро	0.454	0.965	0.825	0.555
	Сигмоидное	0.363	0.827	0.700	0.488
5	Линейное	0.454	0.965	0.825	0.555
	Полиномиальное	0.454	1.000	0.850	0.556
	RBF-ядро	0.454	0.965	0.825	0.555
	Сигмоидное	0.181	0.827	0.650	0.393

Табл. 5. Точность моделей определения прибыльности сельскохозяйственных организаций, производящих овощи открытого грунта, построенных на основании метода опорных векторов с разными ядрами

Tab. 5. Accuracy of models for determining the profitability of agricultural organizations producing open-ground vegetables based on the method of support vectors with different cores

Кол-во влияющих факторов	Применяемое ядро сглаживания	Точность идентификации убыточных организаций	Точность идентификации прибыльных организаций	Общая точность модели	Скорректированная точность модели
1	Линейное	0.000	1.000	0.780	0.347
	Полиномиальное	0.025	1.000	0.786	0.356
	RBF-ядро	0.205	0.985	0.814	0.432
	Сигмоидное	0.205	0.964	0.797	0.431
2	Линейное	0.000	1.000	0.780	0.347
	Полиномиальное	0.256	0.971	0.814	0.456
	RBF-ядро	0.358	0.978	0.842	0.509
	Сигмоидное	0.076	0.978	0.780	0.375
3	Линейное	0.000	1.000	0.780	0.347
	Полиномиальное	0.461	0.978	0.865	0.568
	RBF-ядро	0.487	0.978	0.870	0.584
	Сигмоидное	0.000	0.992	0.775	0.346
4	Линейное	0.025	0.985	0.775	0.355
	Полиномиальное	0.512	0.978	0.876	0.600
	RBF-ядро	0.487	0.978	0.870	0.584
	Сигмоидное	0.076	0.964	0.769	0.374
5	Линейное	0.025	0.985	0.775	0.355
	Полиномиальное	0.256	1.000	0.837	0.457
	RBF-ядро	0.589	1.000	0.910	0.652
	Сигмоидное	0.307	0.863	0.741	0.472

Табл. 6. Точность моделей определения прибыльности сельскохозяйственных организаций, производящих овощи закрытого грунта, построенных на основании метода опорных векторов с разными ядрами

Tab. 6. Accuracy of models for determining the profitability of agricultural organizations producing vegetables of closed ground based on the method of support vectors with different cores

Кол-во влияющих факторов	Применяемое ядро сглаживания	Точность идентификации убыточных организаций	Точность идентификации прибыльных организаций	Общая точность модели	Скорректированная точность модели
1	Линейное	0.000	1.000	0.653	0.311
	Полиномиальное	0.000	1.000	0.653	0.311
	RBF-ядро	0.000	1.000	0.653	0.311
	Сигмоидное	0.000	1.000	0.653	0.311
2	Линейное	0.000	1.000	0.653	0.311
	Полиномиальное	0.000	1.000	0.653	0.311
	RBF-ядро	0.000	1.000	0.653	0.311
	Сигмоидное	0.000	1.000	0.653	0.311
3	Линейное	0.000	1.000	0.653	0.311
	Полиномиальное	0.411	0.906	0.734	0.509
	RBF-ядро	0.000	1.000	0.653	0.311
	Сигмоидное	0.000	1.000	0.653	0.311
4	Линейное	0.000	1.000	0.653	0.311
	Полиномиальное	0.529	0.906	0.775	0.586
	RBF-ядро	0.764	1.000	0.918	0.771
	Сигмоидное	0.000	1.000	0.653	0.311
5	Линейное	0.294	0.875	0.673	0.438
	Полиномиальное	0.529	0.937	0.795	0.588
	RBF-ядро	0.764	1.000	0.918	0.771
	Сигмоидное	0.588	0.812	0.734	0.622

Табл. 7. Точность моделей определения прибыльности сельскохозяйственных организаций, производящих лен, построенных на основании метода опорных векторов с разными ядрами
Tab. 7. Accuracy of models for determining the profitability of agricultural organizations producing flax based on the method of support vectors with different cores

Кол-во влияющих факторов	Применяемое ядро сглаживания	Точность идентификации убыточных организаций	Точность идентификации прибыльных организаций	Общая точность модели	Скорректированная точность модели
1	Линейное	0.886	0.468	0.670	0.498
	Полиномиальное	0.750	0.531	0.637	0.546
	RBF-ядро	0.886	0.510	0.692	0.530
	Сигмоидное	0.727	0.489	0.604	0.510
2	Линейное	0.909	0.446	0.670	0.483
	Полиномиальное	0.863	0.723	0.791	0.716
	RBF-ядро	0.886	0.723	0.802	0.714
	Сигмоидное	0.840	0.489	0.659	0.514
3	Линейное	0.795	0.765	0.780	0.764
	Полиномиальное	0.886	0.872	0.879	0.870
	RBF-ядро	0.840	0.872	0.857	0.838
	Сигмоидное	0.636	0.702	0.670	0.639
4	Линейное	0.818	0.765	0.791	0.762
	Полиномиальное	0.909	0.872	0.890	0.867
	RBF-ядро	0.863	0.872	0.868	0.862
	Сигмоидное	0.795	0.595	0.692	0.601
5	Линейное	0.795	0.744	0.769	0.742
	Полиномиальное	0.909	0.872	0.890	0.867
	RBF-ядро	0.886	0.893	0.890	0.885
	Сигмоидное	0.636	0.638	0.637	0.636

Табл. 8. Точность моделей определения прибыльности сельскохозяйственных организаций, производящих мясо крупного рогатого скота, построенных на основании метода опорных векторов с разными ядрами
 Tab. 8. Accuracy of models for determining the profitability of agricultural organizations producing cattle meat based on the method of support vectors with different cores

Кол-во влияющих факторов	Применяемое ядро сглаживания	Точность идентификации убыточных организаций	Точность идентификации прибыльных организаций	Общая точность модели	Скорректированная точность модели
1	Линейное	0.678	0.549	0.609	0.556
	Полиномиальное	0.000	0.997	0.531	0.262
	RBF-ядро	0.498	0.736	0.625	0.528
	Сигмоидное	0.576	0.629	0.604	0.582
2	Линейное	0.485	0.771	0.638	0.520
	Полиномиальное	0.509	0.742	0.633	0.537
	RBF-ядро	0.521	0.763	0.650	0.547
	Сигмоидное	0.613	0.589	0.600	0.590
3	Линейное	0.558	0.691	0.629	0.572
	Полиномиальное	0.515	0.761	0.646	0.543
	RBF-ядро	0.531	0.761	0.653	0.555
	Сигмоидное	0.468	0.571	0.523	0.486
4	Линейное	0.583	0.677	0.633	0.592
	Полиномиальное	0.554	0.763	0.665	0.574
	RBF-ядро	0.699	0.739	0.720	0.700
	Сигмоидное	0.538	0.599	0.570	0.547
5	Линейное	0.594	0.677	0.638	0.602
	Полиномиальное	0.584	0.745	0.670	0.597
	RBF-ядро	0.692	0.733	0.714	0.693
	Сигмоидное	0.551	0.619	0.587	0.559

Табл. 9. Точность моделей определения прибыльности сельскохозяйственных организаций, производящих мясо свиней, построенных на основании метода опорных векторов с разными ядрами
 Tab. 9. Accuracy of models for determining the profitability of agricultural organizations producing pork based on the method of support vectors with different cores

Кол-во влияющих факторов	Применяемое ядро сглаживания	Точность идентификации убыточных организаций	Точность идентификации прибыльных организаций	Общая точность модели	Скорректированная точность модели
1	Линейное	0.000	1.000	0.529	0.260
	Полиномиальное	0.555	0.697	0.630	0.569
	RBF-ядро	0.651	0.624	0.637	0.625
	Сигмоидное	0.357	0.416	0.388	0.372
2	Линейное	0.311	0.910	0.628	0.411
	Полиномиальное	0.600	0.734	0.671	0.610
	RBF-ядро	0.642	0.669	0.656	0.644
	Сигмоидное	0.655	0.559	0.604	0.564
3	Линейное	0.321	0.906	0.630	0.416
	Полиномиальное	0.605	0.800	0.708	0.617
	RBF-ядро	0.674	0.800	0.740	0.677
	Сигмоидное	0.536	0.595	0.568	0.544
4	Линейное	0.311	0.918	0.632	0.411
	Полиномиальное	0.596	0.787	0.697	0.609
	RBF-ядро	0.683	0.804	0.747	0.686
	Сигмоидное	0.389	0.469	0.431	0.408
5	Линейное	0.371	0.893	0.647	0.447
	Полиномиальное	0.600	0.840	0.727	0.614
	RBF-ядро	0.614	0.816	0.721	0.625
	Сигмоидное	0.642	0.542	0.589	0.549

Табл. 10. Точность моделей определения прибыльности сельскохозяйственных организаций, производящих мясо птицы, построенных на основании метода опорных векторов с разными ядрами
Tab. 10. Accuracy of models for determining the profitability of agricultural organizations producing poultry meat based on the method of support vectors with different cores

Кол-во влияющих факторов	Применяемое ядро сглаживания	Точность идентификации убыточных организаций	Точность идентификации прибыльных организаций	Общая точность модели	Скорректированная точность модели
1	Линейное	0.000	1.000	0.735	0.336
	Полиномиальное	0.043	1.000	0.747	0.353
	RBF-ядро	0.000	1.000	0.735	0.336
	Сигмоидное	0.000	1.000	0.735	0.336
2	Линейное	0.000	1.000	0.735	0.336
	Полиномиальное	0.086	1.000	0.758	0.371
	RBF-ядро	0.000	1.000	0.735	0.336
	Сигмоидное	0.000	1.000	0.735	0.336
3	Линейное	0.000	1.000	0.735	0.336
	Полиномиальное	0.304	0.968	0.793	0.471
	RBF-ядро	0.347	0.953	0.793	0.494
	Сигмоидное	0.000	1.000	0.735	0.336
4	Линейное	0.565	0.968	0.862	0.628
	Полиномиальное	0.695	0.984	0.908	0.724
	RBF-ядро	0.521	0.984	0.862	0.600
	Сигмоидное	0.434	0.968	0.827	0.545
5	Линейное	0.608	0.953	0.862	0.658
	Полиномиальное	0.608	0.968	0.873	0.659
	RBF-ядро	0.608	1.000	0.896	0.660
	Сигмоидное	0.347	0.968	0.804	0.495

Табл. 11. Точность моделей определения прибыльности сельскохозяйственных организаций, производящих молоко, построенных на основании метода опорных векторов с разными ядрами
Tab. 11. Accuracy of models for determining the profitability of agricultural organizations producing milk based on the method of support vectors with different cores

Кол-во влияющих факторов	Применяемое ядро сглаживания	Точность идентификации убыточных организаций	Точность идентификации прибыльных организаций	Общая точность модели	Скорректированная точность модели
1	Линейное	0.669	0.558	0.610	0.564
	Полиномиальное	0.588	0.643	0.618	0.594
	RBF-ядро	0.506	0.726	0.624	0.533
	Сигмоидное	0.762	0.390	0.562	0.432
2	Линейное	0.417	0.812	0.629	0.475
	Полиномиальное	0.493	0.759	0.636	0.527
	RBF-ядро	0.509	0.776	0.652	0.540
	Сигмоидное	0.579	0.626	0.604	0.584
3	Линейное	0.519	0.730	0.632	0.544
	Полиномиальное	0.468	0.778	0.634	0.509
	RBF-ядро	0.578	0.755	0.673	0.594
	Сигмоидное	0.570	0.622	0.598	0.576
4	Линейное	0.520	0.730	0.633	0.545
	Полиномиальное	0.522	0.741	0.639	0.547
	RBF-ядро	0.629	0.757	0.698	0.637
	Сигмоидное	0.569	0.621	0.597	0.575
5	Линейное	0.578	0.669	0.627	0.588
	Полиномиальное	0.541	0.768	0.663	0.564
	RBF-ядро	0.696	0.754	0.727	0.697
	Сигмоидное	0.556	0.617	0.589	0.564

Результаты применения метода SVM с различными функциями представлены в табл. 4.

Соответственно, лучшую точность по обоим критериям показывает модель с полиномиальным ядром и 3–5 значащими факторами, доля правильно классифицированных объектов – 85.0 %.

Рассмотрим организации, производящие овощи открытого грунта (ООГ), с пятью наиболее значащими для них факторами:

- фондовооруженность, тыс. р./чел.;
- полная себестоимость продукции растениеводства, тыс. р.;
- среднегодовая численность работников, чел.;
- мощность 1 единицы техники, л. с./шт.;
- полная себестоимость продукции животноводства, тыс. р.

Результаты применения метода SVM с различными функциями представлены в табл. 5.

Соответственно, лучшую точность по обоим критериям показывает модель с RBF-ядром и пятью значащими факторами, доля правильно классифицированных объектов – 91.0 %.

Рассмотрим организации, производящие овощи закрытого грунта (ОЗГ), с пятью наиболее значащими для них факторами:

- среднемесячная зарплата, р.;
- полная себестоимость продукции растениеводства, тыс. р.;
- среднегодовая численность работников, чел.;
- мощность техники, л. с.;
- мощность 1 единицы техники, л. с./шт.

Результаты применения метода SVM с различными функциями представлены в табл. 6.

Соответственно, лучшую точность по обоим критериям показывает модель с RBF-ядром и 4–5 значащими факторами, доля правильно классифицированных объектов – 91.8 %.

Рассмотрим организации, производящие лен, с пятью наиболее значащими для них факторами:

- стоимость основных средств, тыс. р.;
- общее количество посевных площадей, га;
- всего пашни, га;
- всего сельскохозяйственных угодий, га;
- мощность, приходящаяся на 1 среднегодового работника, л. с./гол.

Результаты применения метода SVM с различными функциями представлены в табл. 7.

Табл. 12. Точность моделей определения прибыльности сельскохозяйственных организаций, производящих яйца, построенных на основании метода опорных векторов с разными ядрами
Tab. 12. Accuracy of models for determining the profitability of agricultural organizations producing eggs based on the method of support vectors with different cores

Кол-во влияющих факторов	Применяемое ядро сглаживания	Точность идентификации убыточных организаций	Точность идентификации прибыльных организаций	Общая точность модели	Скорректированная точность модели
1	Линейное	0.00	1.000	0.736	0.336
	Полиномиальное	0.00	1.000	0.736	0.336
	RBF-ядро	0.00	1.000	0.736	0.336
	Сигмоидное	0.52	0.757	0.694	0.577
2	Линейное	0.00	1.000	0.736	0.336
	Полиномиальное	0.64	0.828	0.778	0.672
	RBF-ядро	0.00	1.000	0.736	0.336
	Сигмоидное	0.00	1.000	0.736	0.336
3	Линейное	0.60	0.928	0.842	0.651
	Полиномиальное	0.68	0.928	0.863	0.710
	RBF-ядро	0.60	0.957	0.863	0.652
	Сигмоидное	0.40	0.842	0.726	0.513
4	Линейное	0.68	0.971	0.894	0.712
	Полиномиальное	0.72	0.971	0.905	0.743
	RBF-ядро	0.68	0.971	0.894	0.712
	Сигмоидное	0.12	0.785	0.610	0.362
5	Линейное	0.68	0.971	0.894	0.712
	Полиномиальное	0.72	0.971	0.905	0.743
	RBF-ядро	0.76	1.000	0.936	0.776
	Сигмоидное	0.16	0.785	0.621	0.380

Соответственно, лучшую точность по обоим критериям показывает модель с RBF-ядром и пятью значащими факторами, доля правильно классифицированных объектов – 89.0 %.

Рассмотрим организации, производящие мясо крупного рогатого скота, с пятью наиболее значащими для них факторами:

- общее количество убранных площадей, га;
- всего пашни, га;

– среднемесячная зарплата, р.;

– мощность техники, л. с.;

– среднегодовая численность работников, чел.

Результаты применения метода SVM с различными функциями представлены в табл. 8.

Соответственно, лучшую точность по обоим критериям показывает модель с RBF-ядром и четырьмя значащими факторами, доля правильно классифицированных объектов – 72.0 %.

Табл. 13. Сводная информация о важности переменных в построенных моделях по методу опорных векторов
 Tab. 13. Summary information about the significance of variables in the constructed models using the support vector method

Фактор	Вид продукции, по которому построена модель											
	Зерно	Рапа	Картофель	Сахарная свекла	Овощи открытого грунта	Овощи закрытого грунта	Лен	Мясо крупного рогатого скота	Мясо свиней	Мясо птицы	Молоко	Яйца
Среднегодовая численность работников	–	3	4	–	3	3	–	5	–	4	–	1
Среднемесячная зарплата	3	–	–	–	–	1	–	3	–	–	5	3
Площадь сельскохозяйственных угодий	–	2	–	5	–	–	4	–	4	1	–	–
Площадь пашни	2	1	–	4	–	–	3	2	2	2	2	–
Условное поголовье скота	4	–	1	–	–	–	–	–	–	–	–	–
Количество сельскохозяйственной техники	–	–	2	–	–	–	–	–	–	–	3	–
Мощность техники	5	–	–	–	–	4	–	4	–	–	–	–
Мощность, приходящаяся на 1 среднегодового работника	–	–	–	–	–	–	5	–	–	5	–	–
Мощность 1 единицы техники	–	–	–	–	4	5	–	–	3	3	–	4
Стоимость основных средств	–	5	3	3	–	–	1	–	–	–	4	2
Стоимость 1 л. с. мощности	–	–	–	–	–	–	–	–	5	–	–	5
Фондовооруженность	–	–	–	–	1	–	–	–	–	–	–	–
Общее количество посевных площадей	–	–	–	–	–	–	2	–	–	–	–	–
Общее количество убранных площадей	1	–	5	1	–	–	–	1	1	–	1	–
Полная себестоимость продукции растениеводства	–	–	–	–	2	2	–	–	–	–	–	–
Полная себестоимость продукции животноводства	–	4	–	2	5	–	–	–	–	–	–	–
Параметры оптимальной модели												
Тип ядра	RBF	Полно-мгальное	RBF	Полно-мгальное	RBF	RBF	RBF	RBF	RBF	Полно-мгальное	RBF	RBF
Количество влияющих факторов	4	5	5	3–5	5	4–5	5	4	4	4	5	5
Точность классификации, %	70.2	68.0	82.0	85.0	91.0	91.8	89.0	72.0	74.7	90.8	72.7	77.6

Рассмотрим организации, производящие мясо свиней, с пятью наиболее значащими для них факторами:

- общее количество убранных площадей, га;
- всего пашни, га;
- мощность 1 единицы техники, л. с./шт.;
- всего сельскохозяйственных угодий, га;
- стоимость 1 л. с. мощности, р.

Результаты применения метода SVM с различными функциями представлены в табл. 9.

Соответственно, лучшую точность по обоим критериям показывает модель с RBF-ядром и четырьмя значащими факторами, доля правильно классифицированных объектов – 74.7 %.

Рассмотрим организации, производящие мясо птицы, с пятью наиболее значащими для них факторами:

- всего сельскохозяйственных угодий, га;
- всего пашни, га;
- мощность 1 единицы техники, л. с./шт.;
- среднегодовая численность работников, чел.;
- приходящаяся мощность на 1 среднегодового работника, л. с./гол.

Результаты применения метода SVM с различными функциями представлены в табл. 10.

Соответственно, лучшую точность по обоим критериям показывает модель с полиномиальным ядром и четырьмя значащими факторами, доля правильно классифицированных объектов – 90.8 %.

Рассмотрим организации, производящие молоко, с пятью наиболее значащими для них факторами:

- общее количество убранных площадей, га;
- всего пашни, га;
- мощность техники, л. с.;
- стоимость основных средств, тыс. р.;
- среднемесячная зарплата, р.

Результаты применения метода SVM с различными функциями представлены в табл. 11.

Соответственно, лучшую точность по обоим критериям показывает модель с RBF-ядром и пятью значащими факторами, доля правильно классифицированных объектов – 72.7 %.

Рассмотрим организации, производящие яйца, с пятью наиболее значащими для них факторами:

- среднегодовая численность работников, чел.;
- стоимость основных средств, тыс. р.;
- среднемесячная зарплата, р.;
- мощность 1 единицы техники, л. с./шт.;
- стоимость 1 л. с. мощности, р.

Результаты применения метода SVM с различными функциями представлены в табл. 12.

Соответственно, лучшую точность по обоим критериям показывает модель с RBF-ядром и пятью значащими факторами, доля правильно классифицированных объектов – 77.6 %.

Обсуждение результатов. Обобщим данные, представленные в табл. 13. В таблице представлены ранги пяти самых значимых переменных для каждой из построенных моделей.

Анализ сводной информации позволяет сделать вывод, что наиболее важными факторами, влияющими на эффективность предсказания убыточности/прибыльности сельскохозяйственных организаций с помощью метода опорных векторов, являются ресурсные показатели, показывающие количество работников, размер пашни, стоимость основных средств. Наиболее эффективен алгоритм при применении RBF-ядра. Для организаций, производящих ту или иную продукцию животноводства, его точность значительно ниже, чем для организаций, производящих ту или иную продукцию растениеводства.

Список литературы

1. Kouziokas G. N. A new W-SVM kernel combining PSO-neural network transformed vector and Bayesian optimized SVM in GDP forecasting // *Engin. Appl. of Artificial Intelligence*. 2020. № 92. P. 103650. doi: 10.1016/j.engappai.2020.103650.
2. Kong R., Zhang B. Autocorrelation kernel functions for support vector machines // *Third Intern. Conf. on Natural Computation (ICNC 2007)*. Haikou, China. 2007. P. 512–516. doi: 10.1109/ICNC.2007.276.
3. Moghaddam V. H., Hamidzadeh J. New Hermite orthogonal polynomial kernel and combined kernels in support vector machine classifier // *Pattern Recognition*. 2016. № 60. P. 921–935. doi: 10.1016/j.patcog.2016.07.004.
4. Lin Z., Yan L. A support vector machine classifier based on a new kernel function model for hyperspectral data // *GIScience Remote Sens*. 2015. № 53. P. 85–101. doi: 10.1080/15481603.2015.1114199.
5. Tran H. D., Li H. Probabilistic distance SVM with Hellinger-exponential kernel for sound event classification // *Intern. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)*. Prague, Czech Republic. 2011. P. 2272–2275. doi: 10.1109/ICASSP.2011.5946935.
6. The mixed kernel function SVM-based point cloud classification / C. Chen, X. Li, A. N. Belkacem, Z. Qiao, E. Dong, W. Tan, D. Shin // *Int. J. Precis. Eng. Manuf*. 2019. № 20. P. 737–747. doi: 10.1007/s12541-019-00102-3.
7. A novel incremental kernel nonparametric SVM model (iKN-SVM) for data classification: An application to face detection / A. Soula, K. Tbarki, R. Ksantini, S. B. Said, Z. Lachiri // *Eng. Appl. Artif. Intell*. 2020. № 89. P. 103468. doi: 10.1016/j.engappai.2019.103468.

8. Zeng J., Qiao W. Short-term wind power prediction using a wavelet support vector machine // *IEEE Trans. Sustain. Energy*. 2012. Vol. 3, No. 2. P. 255–264. doi: 10.1109/TSSTE.2011.2180029.
9. Aurangzeb K., Ayub N., Alhussein M. Aspect based multi-labeling using SVM based ensembler // *IEEE ACCESS*. Vol. 9. P. 26026–26040. doi: 10.1109/ACCESS.2021.3055768.
10. Annual electric load forecasting by a least squares support vector machine with a fruit fly optimization algorithm / H. Li, S. Guo, H. Zhao, Ch. Su, B. Wanget // *Energies*. 2012. Vol. 5, No. 11. P. 4430–4445. doi: 10.3390/en 5114430.
11. Research on and application of diminishing step fruit fly optimization algorithm / N. Jianping, W. Bing, L. Hongru, B. Xu // *J. of Shenzhen University: Sci. and Engin.* 2014. № 31(4). P. 367–373. doi: 10.3724/SP.J.1249.2014.04367.
12. Cao G., L. Wu. Support vector regression with fruit fly optimization algorithm for seasonal electricity consumption forecasting // *Energy*. 2016. № 115. P. 734–745. doi: 10.1016/j.energy.2016.09.065.
13. Ye F., Lou X. Y., Sun L. F. An improved chaotic fruit fly optimization based on a mutation strategy for simultaneous feature selection and parameter optimization for SVM and its applications // *Plos One*. 2017. № 12(4). P. 1–36. doi: 10.1371/journal.pone.0173516.
14. A novel F-SVM based on FOA for improving SVM performance / Q. Gu, Y. Chang, X. Li, Z. Chang, Z. Feng // *Expert Systems with Applications*. 2020. № 165(10). P. 113713. doi: 10.1016/j.eswa.2020.113713.
15. Evolutionary wrapper approaches for training set selection as preprocessing mechanism for support vector machines: Experimental evaluation and support vector analysis / N. Verbiest, J. Derrac, C. Cornelis, S. Garcia, F. Herrera // *Appl. Soft Computing*. 2016. № 38. P. 10–22. doi: 10.1016/j.asoc.2015.09.006.
16. Data selection based on decision tree for SVM classification on large data sets / J. Cervantes, F. G. Lamont, A. Lopez-Chau, L. R. Mazahua, J. S. Ruiz // *Appl. Soft Computing*. 2015. № 37. P. 787–798. doi: 10.1016/j.asoc.2015.08.048.
17. Multi-objective optimization and meta-learning for SVM parameter selection / P. B. Miranda, R. B. Prudencio, A. C. P. de Carvalho, C. Soares // *The 2012 Intern. Joint Conf. on Neural Networks (IJCNN)*. Brisbane, Australia. 2012. P. 1–8. doi: 10.1109/IJCNN.2012.6252378.
18. Surrogate-assisted multiobjective model selection for support vector machines / A. Rosales-Perez, J. A. Gonzalez, C. A. C. Coello, H. J. Escalante, C. A. Reyes-Garcia // *Neurocomputing*. 2015. № 150. P. 163–172. doi: 10.1016/j.neucom.2014.08.075.
19. Jung H. G., Kim G. Support vector number reduction: Survey and experimental evaluations // *IEEE Transactions on Intelligent Transportation Systems*. 2013. № 15(2). P. 463–476. doi: 10.1109/TITS.2013.2282635.
20. Nie F., Zhu W., Li X. Decision Tree SVM: An extension of linear SVM for non-linear classification // *Neurocomputing*. 2015. № 401(3). doi: 10.1016/j.neucom.2019.10.051.
21. Lin C. F., Wang S. D. Fuzzy support vector machines // *IEEE Trans. Neural Netw.* 2002. № 13(2). P. 464–471. doi: 10.1109/72.991432.
22. Fuzzy quasi-linear SVM classifier: Design and analysis / C. Yang, S.-K. Oh, B. Yang, W. Pedrycz // *Fuzzy Sets and Systems*. 2020. № 413(1). doi: 10.1016/j.fss.2020.05.010.
23. SVM-Boosting based on Markov resampling: Theory and algorithm / H. Jiang, B. Zou, C. Xu, J. Xu, Y. Y. Tang // *Neural Networks*. 2020. № 131. P. 276–290. doi: 10.1016/j.neunet.2020.07.036.
24. Zhang X., Li A., Pan R. Stock trend prediction based on a new status box method and adaboost probabilistic support vector machine // *Appl. Soft Comput.* 2016. № 49. P. 385–398. doi: 10.1016/j.asoc.2016.08.026.
25. Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval / D. Tao, X. Tang, X. Li, X. Wu // *IEEE Trans. Pattern Anal. Mach. Intell.* 2006. № 28(7). P. 1088–1099. doi: 10.1109/TRAMI.2006.134.
26. Mordelet F., Vert J.-P. A bagging SVM to learn from positive and unlabeled examples // *Pattern Recognit. Lett.* 2014. № 37. P. 201–209. doi: 10.1016/j.patrec.2013.06.010.
27. Combined TBATS and SVM model of minimum and maximum air temperatures applied to wheat yield prediction at different locations in Europe / M. Gos, J. Krzyszczyk, P. Baranowski, M. Murat, I. Malinowska // *Agricultural and Forest Meteorology*. 2020. № 281. 107827.
28. Application of an SVM-based regression model to the air quality study at local scale in the Avilés urban area (Spain) / A. S. Sánchez, P. G. Nieto, P. R. Fernández, J. del Coz Díaz, F. J. Iglesias-Rodríguez // *Mathematical and Computer Modelling*. 2011. № 54. P. 1453–1466. doi: 10.1016/j.mcm.2011.04.017.
29. Budi A. S., Susilo P. H., Nafiyah N. SVM algorithm for predicting rice yields // *J. Teknologi Informasi dan Pendidikan*. 2020. Vol. 13, no. 2. P. 50–54. doi: 10.24036/tip.v13i2.341.
30. Mardhika D. A., Setiawan B. D., Wihandika R. C. Penerapan algoritma support vector regression pada peramalan hasil panen padi studi kasus kabupaten malang // *J. Pengembangan Teknologi Informasi dan Ilmu Komputer*. 2019. Vol. 3, no. 10. P. 9402–9412. doi: 10.47065/bits.v3i2.1004.
31. The extent that certain dairy farmer attitudes and behaviors are associated with farm business profitability / N. W. O'Leary, R. M. Bennett, R. B. Tranter, P. J. Jones // *J. Dairy Sci.* 2017. № 101(12). P. 11275–11284. doi: 10.3168/jds.2017-14307.
32. Kocsis J., Major K. A. General overview of agriculture and profitability in agricultural enterprises in Central Europe // *Managing Agricultural Enterprises* / ed. P. Bryła. Springer International Publishing. 2018. P. 243–265. doi: 10.1007/978-3-319-59891-8_12.
33. Анализ результативности метода опорных векторов при статистической обработке больших данных / А. Д. Черемухин, А. А. Шамин, М. О. Колбанев, В. В. Цехановский // *Изв. СПбГЭТУ «ЛЭТИ»*. 2021. № 2. С. 58–68.

Информация об авторах

Черемухин Артем Дмитриевич – канд. экон. наук, доцент кафедры «Математика и вычислительная техника», Нижегородский государственный инженерно-экономический университет, ул. Октябрьская, д. 22а, Княгинино, Нижегородская область, 606340, Россия.

E-mail: ngieu.cheremuhin@yandex.ru

<https://orcid.org/0000-0003-4076-5916>

Шамин Алексей Анатольевич – канд. экон. наук, директор Института информационных технологий и систем связи, Нижегородский государственный инженерно-экономический университет, ул. Октябрьская, д. 22а, Княгинино, Нижегородская область, 606340, Россия.

E-mail: ngiei-spo@mail.ru

<https://orcid.org/0000-0003-4138-6256>

Колбанев Михаил Олегович – д-р техн. наук, профессор кафедры информационных систем СПбГЭТУ «ЛЭТИ».

E-mail: mokolbanev@mail.ru

<https://orcid.org/0000-0003-4825-6972>

Цехановский Владислав Владимирович – канд. техн. наук, доцент, профессор кафедры информационных систем СПбГЭТУ «ЛЭТИ».

E-mail: vvcehanovsky@mail.ru

<https://orcid.org/0000-0003-0287-1806>

References

1. Kouziokas G. N. A new W-SVM kernel combining PSO-neural network transformed vector and Bayesian optimized SVM in GDP forecasting // *Engin. Appl. of Artificial Intelligence*. 2020. № 92. P. 103650. doi: 10.1016/j.engappai.2020.103650.
2. Kong R., Zhang B. Autocorrelation kernel functions for support vector machines // *Third Intern. Conf. on Natural Computation (ICNC 2007)*. Haikou, China. 2007. P. 512–516. doi: 10.1109/ICNC.2007.276.
3. Moghaddam V. H., Hamidzadeh J. New Hermite orthogonal polynomial kernel and combined kernels in support vector machine classifier // *Pattern Recognition*. 2016. № 60. P. 921–935. doi: 10.1016/j.patcog.2016.07.004.
4. Lin Z., Yan L. A support vector machine classifier based on a new kernel function model for hyperspectral data // *GIScience Remote Sens*. 2015. № 53. P. 85–101. doi: 10.1080/15481603.2015.1114199.
5. Tran H. D., Li H. Probabilistic distance SVM with Hellinger-exponential kernel for sound event classification // *Intern. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)*. Prague, Czech Republic. 2011. P. 2272–2275. doi: 10.1109/ICASSP.2011.5946935.
6. The mixed kernel function SVM-based point cloud classification / C. Chen, X. Li, A. N. Belkacem, Z. Qiao, E. Dong, W. Tan, D. Shin // *Int. J. Precis. Eng. Manuf*. 2019. № 20. P. 737–747. doi: 10.1007/s12541-019-00102-3.
7. A novel incremental kernel nonparametric SVM model (iKN-SVM) for data classification: An application to face detection / A. Soula, K. Tbariki, R. Ksantini, S. B. Said, Z. Lachiri // *Eng. Appl. Artif. Intell*. 2020. № 89. P. 103468. doi: 10.1016/j.engappai.2019.103468.
8. Zeng J., Qiao W. Short-term wind power prediction using a wavelet support vector machine // *IEEE Trans. Sustain. Energy*. 2012. Vol. 3, No. 2. P. 255–264. doi: 10.1109/TSTE.2011.2180029.
9. Aurangzeb K., Ayub N., Alhussein M. Aspect based multi-labeling using SVM based ensembler // *IEEE ACCESS*. Vol. 9. P. 26026–26040. doi: 10.1109/ACCESS.2021.3055768.
10. Annual electric load forecasting by a least squares support vector machine with a fruit fly optimization algorithm / H. Li, S. Guo, H. Zhao, Ch. Su, B. Wang // *Energies*. 2012. Vol. 5, No. 11. P. 4430–4445. doi: 10.3390/en5114430.
11. Research on and application of diminishing step fruit fly optimization algorithm / N. Jianping, W. Bing, L. Hongru, B. Xu // *J. of Shenzhen University: Sci. and Engin*. 2014. № 31(4). P. 367–373. doi: 10.3724/SP.J.1249.2014.04367.
12. Cao G., L. Wu. Support vector regression with fruit fly optimization algorithm for seasonal electricity consumption forecasting // *Energy*. 2016. № 115. P. 734–745. doi: 10.1016/j.energy.2016.09.065.
13. Ye F., Lou X. Y., Sun L. F. An improved chaotic fruit fly optimization based on a mutation strategy for simultaneous feature selection and parameter optimization for SVM and its applications // *Plos One*. 2017. № 12(4). P. 1–36. doi: 10.1371/journal.pone.0173516.
14. A novel F-SVM based on FOA for improving SVM performance / Q. Gu, Y. Chang, X. Li, Z. Chang, Z. Feng // *Expert Systems with Applications*. 2020. № 165(10). P. 113713. doi: 10.1016/j.eswa.2020.113713.

15. Evolutionary wrapper approaches for training set selection as preprocessing mechanism for support vector machines: Experimental evaluation and support vector analysis / N. Verbiest, J. Derrac, C. Cornelis, S. Garcia, F. Herrera // *Appl. Soft Computing*. 2016. № 38. P. 10–22. doi: 10.1016/j.asoc.2015.09.006.
16. Data selection based on decision tree for SVM classification on large data sets / J. Cervantes, F. G. Lamont, A. Lopez-Chau, L. R. Mazahua, J. S. Ruiz // *Appl. Soft Computing*. 2015. № 37. P. 787–798. doi: 10.1016/j.asoc.2015.08.048.
17. Multi-objective optimization and meta-learning for SVM parameter selection / P. B. Miranda, R. B. Prudencio, A. C. P. de Carvalho, C. Soares // *The 2012 Intern. Joint Conf. on Neural Networks (IJCNN)*. Brisbane, Australia. 2012. P. 1–8. doi: 10.1109/IJCNN.2012.6252378.
18. Surrogate-assisted multiobjective model selection for support vector machines / A. Rosales-Perez, J. A. Gonzalez, C. A. C. Coello, H. J. Escalante, C. A. Reyes-Garcia // *Neurocomputing*. 2015. № 150. P. 163–172. doi: 10.1016/j.neucom.2014.08.075.
19. Jung H. G., Kim G. Support vector number reduction: Survey and experimental evaluations // *IEEE Trans. on Intelligent Transp. Systems*. 2013. № 15(2). P. 463–476. doi: 10.1109/TITS.2013.2282635.
20. Nie F., Zhu W., Li X. Decision Tree SVM: An extension of linear SVM for non-linear classification // *Neurocomputing*. 2015. № 401(3). doi: 10.1016/j.neucom.2019.10.051.
21. Lin C. F., Wang S. D. Fuzzy support vector machines // *IEEE Trans. Neural Netw.* 2002. № 13(2). P. 464–471. doi: 10.1109/72.991432.
22. Fuzzy quasi-linear SVM classifier: Design and analysis / C. Yang, S.-K. Oh, B. Yang, W. Pedrycz // *Fuzzy Sets and Systems*. 2020. № 413(1). doi: 10.1016/j.fss.2020.05.010.
23. SVM-Boosting based on Markov resampling: Theory and algorithm / H. Jiang, B. Zou, C. Xu, J. Xu, Y. Y. Tang // *Neural Netw.* 2020. № 131. P. 276–290. doi: 10.1016/j.neunet.2020.07.036.
24. Zhang X., Li A., Pan R. Stock trend prediction based on a new status box method and adaboost probabilistic support vector machine // *Appl. Soft Comput.* 2016. № 49. P. 385–398. doi: 10.1016/j.asoc.2016.08.026.
25. Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval / D. Tao, X. Tang, X. Li, X. Wu // *IEEE Trans. Pattern Anal. Mach. Intell.* 2006. № 28(7). P. 1088–1099. doi: 10.1109/TRAMI.2006.134.
26. Mordelet F., Vert J.-P. A bagging SVM to learn from positive and unlabeled examples // *Pattern Recognit. Lett.* 2014. № 37. P. 201–209. doi: 10.1016/j.patrec.2013.06.010.
27. Combined TBATS and SVM model of minimum and maximum air temperatures applied to wheat yield prediction at different locations in Europe / M. Gos, J. Krzyszczak, P. Baranowski, M. Murat, I. Malinowska // *Agricultural and Forest Meteorology*. 2020. № 281. 107827.
28. Application of an SVM-based regression model to the air quality study at local scale in the Avilés urban area (Spain) / A. S. Sánchez, P. G. Nieto, P. R. Fernández, J. del Coz Díaz, F. J. Iglesias-Rodríguez // *Math. and Comp. Modelling*. 2011. № 54. P. 1453–1466. doi: 10.1016/j.mcm.2011.04.017.
29. Budi A. S., Susilo P. H., Nafiyah N. SVM algorithm for predicting rice yields // *J. Teknologi Informasi dan Pendidikan*. 2020. Vol. 13, no. 2. P. 50–54. doi: 10.24036/tip.v13i2.341.
30. Mardhika D. A., Setiawan B. D., Wihandika R. C. Penerapan algoritma support vector regression pada peramalan hasil panen padi studi kasus kabupaten malang // *J. Pengembangan Teknologi Informasi dan Ilmu Komputer*. 2019. Vol. 3, no. 10. P. 9402–9412. doi: 10.47065/bits.v3i2.1004.
31. The extent that certain dairy farmer attitudes and behaviors are associated with farm business profitability / N. W. O'Leary, R. M. Bennett, R. B. Tranter, P. J. Jones // *J. Dairy Sci.* 2017. № 101(12). P. 11275–11284. doi: 10.3168/jds.2017-14307.
32. Kocsis J., Major K. A. General Overview of Agriculture and Profitability in Agricultural Enterprises in Central Europe // *Managing Agricultural Enterprises* / ed. P. Bryła. Springer International Publishing. 2018. P. 243–265. doi: 10.1007/978-3-319-59891-8_12.
33. Analiz rezul'tativnosti metoda opornyh vektorov pri statisticheskoy obrabotke bol'shih dannyh / A. D. Chere-muhin, A. A. SHamin, M. O. Kolbanev, V. V. Cehanovskij // *Izv. SPbGETU «LETI»*. 2021. № 2. S. 58–68. (In Russ.).

Information about the authors

Artem D. Cheremuhin – Cand. Sci. (Econ.), Associate Professor of the Department of Mathematics and Computer Engineering, Nizhny Novgorod Engineering-Economic State University, Oktyabrskaya str., 22a, Knyaginino, Nizhny Novgorod region, 606340, Russia.

E-mail: ngieu.cheremuhin@yandex.ru
<https://orcid.org/0000-0003-4076-5916>

Alexey A. Shamin – Cand. Sci. (Econ.), Director of the Institute of Information Technologies and Communication Systems, Nizhny Novgorod Engineering-Economic State University, Oktyabrskaya str., 22a, Knyaginino, Nizhny Novgorod region, 606340, Russia.

E-mail: ngiei-spo@mail.ru
<https://orcid.org/0000-0003-4138-6256>

Mikhail O. Kolbanev – Dr Sci. (Eng.), Professor of the Department of Information Systems, Saint Petersburg Electrotechnical University.

E-mail: mokolbanev@mail.ru

<https://orcid.org/0000-0003-4825-6972>

Vladislav V. Tsehanovsky – Cand. Sci. (Eng.), Associate Professor, Professor of the Department of Information Systems, Saint Petersburg Electrotechnical University.

E-mail: vvcehanovsky@mail.ru

<https://orcid.org/0000-0003-0287-1806>

Статья поступила в редакцию 17.12.2022; принята к публикации после рецензирования 13.01.2023; опубликована онлайн 25.04.2023.

Submitted 17.12.2022; accepted 13.01.2023; published online 25.04.2023.
