

УДК 004.032:004.75

АНСАМБЛИ НЕЙРОННЫХ СЕТЕЙ ДЛЯ КЛАССИФИКАЦИИ СОСТОЯНИЙ И РЕЖИМОВ ФУНКЦИОНИРОВАНИЯ ПОДСИСТЕМ КОСМИЧЕСКОГО АППАРАТА



М.М. Лукашевич
Доцент кафедры информационных систем управления БГУ, кандидат технических наук, доцент
LukashevichMM@bsu.by



В.Ю. Скобцов
Доцент кафедры компьютерных технологий и программной инженерии Санкт-Петербургского государственного университета аэрокосмического приборостроения, кандидат технических наук, доцент
vasko_vasko@mail.ru
и



А.В. Инютин
Зав. лабораторией №222 ОИПИ,
avin@lsi.bas-net.by



В.В. Ганченко
ст. науч. сотр. лаборатории №222 ОИПИ, кандидат технических наук,
ganchenko@lsi.bas-net.by

М.М. Лукашевич

Окончила учреждение образования «Белорусский государственный университет информатики и радиоэлектроники». Область научных интересов связана с разработкой методов и алгоритмов машинного обучения для решения задач анализа данных и компьютерного зрения.

В.Ю. Скобцов

Окончил Донецкий государственный университет. Область научных интересов связана с разработкой и исследованием методов интеллектуального анализа данных, машинного обучения, мягких вычислений, логико-вероятностного моделирования и их применением в области анализа данных функционирования информационных и аппаратных систем, их надежности и живучести, диагностики программно-аппаратных систем.

А.В. Инютин

Окончил Белорусский государственный университет транспорта. Заведующий лабораторией идентификации систем ОИПИ НАН Беларуси. Проводит научные исследования в области обработки и распознавания изображений методами математической морфологии, обработки данных ДЗЗ и телеметрии космических аппаратов.

В.В. Ганченко

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с применением компьютерного зрения и машинного обучения в задачах обработки изображений.

Аннотация. Выполнен разведочный анализ и подготовка наборов данных для решения задач классификации состояний подсистемы КА. Рассмотрены и реализованы ансамблевые методы машинного обучения для решения задач классификации состояния подсистем космического аппарата (КА). Полученные результаты носят прикладной характер и являются элементом комплексной автоматизации и интеллектуализации процессов оценивания и анализа данных телеметрической информации МКА (малого космического аппарата).

Ключевые слова: нейронные сети, ансамблевые методы, классификация, бортовая аппаратура (БА), телеметрическая информация (ТМИ), малый космический аппарат (МКА).

Введение. Одной из важнейших задач на всех этапах жизненного цикла малого космического аппарата (МКА) является анализ их телеметрической информации о функционировании бортовой аппаратуры (БА) МКА с точки зрения определения их технического состояния для обеспечения их штатного и безопасного функционирования. Актуальность обусловлена прежде всего тем, что одной из основных причин потерь МКА являются отказы, сбои и некорректная работа систем БА МКА. Большое количество информации, поступающей и накапливающейся в специализированных банках данных о БА МКА, может быть эффективно использовано для совершенствования процесса определения технического состояния МКА. В связи с тем, что данные о функционировании БА МКА, включающие телеметрические измерения, представляют собой разнородные нерегулярные многомерные данные, актуальным является исследование, разработка и применение моделей, которые позволяют анализировать такого рода данные с возможностью извлечения из них полезной информации и последующего построения с их использованием классификационных и прогностических моделей с целью определения технического состояния МКА для принятия корректных управляющих и эксплуатационных решений в процессе функционирования МКА. Методы машинного обучения, искусственного интеллекта и математической статистики являются на текущий момент одними из самых перспективных и широко используемых подходов в анализе данных высокотехнологичных систем. В современных условиях для решения перечисленных задач, обеспечения требуемой степени автономности, качества и оперативности управления такими сложными объектами как МКА необходимо выполнить комплексную автоматизацию и интеллектуализацию процессов оценивания и многомодельного анализа данных телеметрической информации МКА. Однако, в большинстве случаев на практике автоматизация выполнена, в лучшем случае, лишь частично, и многое делается зачастую вручную, базируясь на эвристических правилах [1]. Таким образом, задача интеллектуального анализа данных телеметрии БА МКА с целью определения технического состояния БА МКА является актуальной и востребованной [1, 2].

Целью исследования являлась разработка моделей машинного обучения для классификации состояния и режимов функционирования подсистем космического аппарата (КА).

Разведочный анализ данных. В качестве данных для построения моделей машинного обучения были использованы данные телеметрической информации (ТМИ) звездного датчика (ЗД) БА Белорусского космического аппарата (БКА) за 2015, 2021 годы. Перечень признаков ТМИ ЗД БКА приведен в таблице 1.

Таблица 1. Перечень телеметрической информации ЗД БКА

Название датчика	Значение датчика
УАС	координаты по солнцу
УВС	координаты по солнцу
ТОК58	ток на ЗД
STRQ1-1	значения кватернионов
STRQ2-1	значения кватернионов
STRQ3-1	значения кватернионов
STRQ4-1	значения кватернионов
TSTR1DSP	температура процессора обработки сигнала
TSTR1CCD	температура ПЗС матрицы (которая делает снимки)
Class1	0 – нештатного состояния подсистемы 1 – штатного состояния подсистемы
Class	0 – завершено 2 – инициализация разворота 3 – ожидание разворота 4 – разворот 8 – отмена

Данные сохранены в файл в формате .csv, всего 2 384 627 объектов, описанных 11 признаками (столбцами). Два столбца («Class1», «Class») являются целевыми или метками классов для 2-х классовых задача и для 5-ти классовых задачи классификации состояния подсистемы. Для набора данных выполнен разведочный анализ данных (EDA, *exploration data analysis*) с использованием библиотеки pandas. Результаты вывода первых 5 строк представлены на рисунке 1.

	УВС	УАС	ТОК58	STRQ1	STRQ2	STRQ3	STRQ4	TSTR1CCD	TSTR1DSP	Class	Class1
0	0.000	0.000	139.0	0.137	-0.312	0.78	0.524	9.849	23.41	0	0
1	0.000	0.000	135.0	0.137	-0.312	0.78	0.524	9.849	23.41	0	0
2	0.000	0.000	135.0	0.132	-0.315	0.78	0.525	9.849	23.41	0	0
3	46.865	32.324	135.0	0.132	-0.315	0.78	0.525	9.849	23.41	0	0
4	21.464	19.704	135.0	0.132	-0.315	0.78	0.525	9.849	23.41	0	0

Рисунок 1. Вывод первых 5 строк таблицы данных

Общая информация о наборе данных, тип каждого признака, а также наличие или отсутствие пропусков представлены на рисунке 2. Статистическая информация по каждому числовому признаку представлена на рисунке 3. Визуализация матрицы корреляции признаков представлена на рисунке 4.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2384627 entries, 0 to 2384626
Data columns (total 11 columns):
#   Column      Dtype
---  ---
0   YBC         float64
1   YAC         float64
2   TOK58       float64
3   STRQ1       float64
4   STRQ2       float64
5   STRQ3       float64
6   STRQ4       float64
7   TSTR1CCD    float64
8   TSTR1DSP    float64
9   Class       int64
10  Class1      int64
dtypes: float64(9), int64(2)
memory usage: 200.1 MB
```

Рисунок 2. Общая информация о наборе данных, тип каждого признака, а также наличие или отсутствие пропусков

	YBC	YAC	TOK58	STRQ1	STRQ2	STRQ3	STRQ4	TSTR1CCD	TSTR1DSP	Class	Class1
count	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06
mean	-1.631350e+01	1.201425e+01	8.774924e+01	2.154371e-01	2.675968e-02	3.409743e-01	4.153858e-01	1.090822e+01	2.866912e+01	7.521285e-01	1.138807e-01
std	4.555501e+01	1.849339e+01	1.968790e+02	3.241060e-01	4.108809e-01	4.478304e-01	4.317669e-01	3.391021e+03	2.563074e+03	2.146762e+00	3.176664e-01
min	-9.000000e+00	-9.000000e+01	0.000000e+00	-7.050000e-01	-7.060000e-01	-7.030000e-01	-7.040000e-01	-4.121000e+01	-4.121000e+01	0.000000e+00	0.000000e+00
25%	-5.247600e+01	3.594000e+00	0.000000e+00	9.000000e-03	-3.160000e-01	-5.000000e-03	1.310000e-01	3.064000e+00	2.341000e+01	0.000000e+00	0.000000e+00
50%	-2.725700e+01	1.100800e+01	1.330000e+02	2.550000e-01	-9.100000e-02	4.410000e-01	5.370000e-01	6.969000e+00	2.471400e+01	0.000000e+00	0.000000e+00
75%	1.650900e+01	1.970000e+01	1.380000e+02	3.990000e-01	3.250000e-01	7.330000e-01	7.810000e-01	1.102200e+01	2.669400e+01	0.000000e+00	0.000000e+00
max	9.000000e+01	9.000000e+01	6.543300e+04	9.990000e-01	1.000000e+00	1.000000e+00	1.000000e+00	3.177054e+06	2.296978e+06	8.000000e+00	1.000000e+00

Рисунок 3. Статистическая информация по каждому числовому признаку

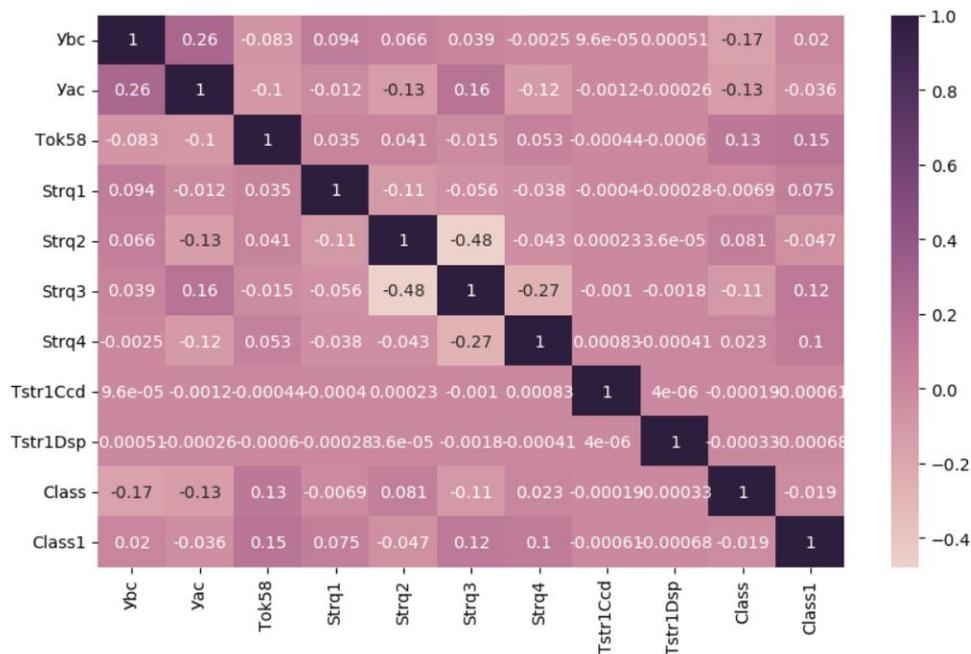


Рисунок 4. Визуализация матрицы корреляции признаков

Описание целевых признаков: а) уникальные значения целевых признаков, число объектов с уникальными значениями для целевых признаков представлено на рисунке 5.

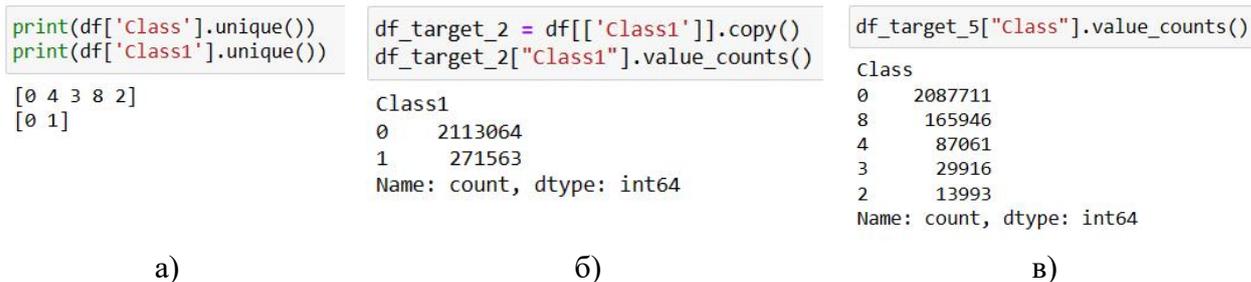


Рисунок 5. Описание целевых признаков: а) уникальные значения целевых признаков, б) и в) число объектов с уникальными значениями для целевых признаков

В результате разведочного анализа данных установлено, что пропущенные значения признаков отсутствуют, все признаки являются числовыми, набор данных с учётом меток классов («Class», «Class1») является несбалансированными. Метки «Class» для дальнейшего этапа построения классификаторов требуют перекодировки значений, таблица 2.

Таблица 2. Перекодировка меток для «Class»

«Class»	Режим подсистемы ориентации и стабилизации	Перекодировка классов	Число объектов
0	Завершено	0	2 087 711
8	Отмена	1	165 946
4	Разворот	2	87 061
3	Ожидание разворота	3	29 916
2	Инициализация разворота	4	13 993

Балансировка данных. В задаче классификации данные называются несбалансированными (*Imbalanced Data*), если в обучающей выборке доли объектов разных классов существенно различаются, также говорят, что классы не сбалансированы. Для решения существующей проблемы было использовано два подхода: замена большого класса подвыборкой по мощности равной малому классу (недосэмплированием, *Undersampling the majority class*) и увеличение в размерах малого класса (пересемплирование, *Oversampling the minority class*). Простейшая стратегия недосэмплирования – взять случайную подвыборку, простейшая стратегия пересэмплирования – продублировать объекты малого класса [3-5], рисунок 10.

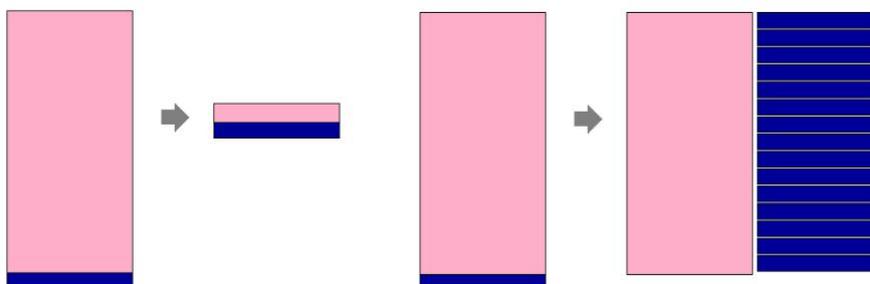


Рисунок 6. Undersampling (слева) и Oversampling (справа)

Подготовка данных для обучения ансамблей нейронных сетей включала в себя подготовку двух набора для 2-классовой задачи классификации и 5-ти классовой задача классификации с назначением столбцов «Class1» и «Class» целевыми столбцами соответственно («Набор данных 1» и «Набор данных 2»). Далее выполнялось разбиение

наборов данных на обучающую и тестовую выборки в соотношении 70% и 30% соответственно. Балансировка данных выполнялась только для обучающих наборов, чтобы получить независимые оценки точности классификации на тестовых наборах и избежать «утечки данных». Недосэмплирование и пересемплирование выполнялось средствами библиотеки *sklearn*. В случае «Набора данных 1» выполнялось недосемплирование для класса «0» и пересемплирование для класса «1». Статистика по исходному «Набору данных 1» представлена в таблице 3, результаты разбиение на обучающую и тестовую выборки - в таблице 4, результаты балансировки – в таблице 5.

Таблица 3. Исходных «Набор данных 1» (2 класса)

Класс	Число объектов
0	2 113 064
1	271 563

Таблица 4. Результаты разбиение на обучающую и тестовую выборки «Набора данных 1»

Набор данных, число объектов	Обучающая выборка, число объектов	Тестовая выборка, число объектов
2 384 627	1 669 238	715 389
0 класс – 2 113 064 1 класс – 271 563	0 класс – 1 479 190 1 класс – 190 048	0 класс – 633 874 1 класс - 81 515

Таблица 5. Результаты балансировки обучающей выборки «Набора данных 1»

Класс	Число объектов после балансировки	Использованная техника
0 класс	190 094	уменьшено число образцов
1 класс	190 094	увеличено число образцов

Для «Набора данных 2» выполнялось недосемплирование для класса «0» и пересемплирование для классов «1» и «2». Статистика по исходному «Набору данных 2» представлена в таблице 6, результаты разбиение на обучающую и тестовую выборки – в таблице 7, результаты балансировки – в таблице 8.

Таблица 6. Исходных «Набор данных 2» (5 класса)

Класс	Число объектов
0	1 461 397
1	9 795
2	20 940
3	60 943
4	116 162

Таблица 7. Результаты разбиение на обучающую и тестовую выборки «Набора данных 2»

Набор данных, число объектов	Обучающая выборка, число объектов	Тестовая выборка, число объектов
2 384 627	1 669 238	715 389
0 класс – 2 087 711 1 класс – 13 993 2 класс – 29 916 3 класс – 87 061 4 класс – 165 946	0 класс – 1 461 397 1 класс – 9 795 2 класс – 20 941 3 класс – 60 943 4 класс – 116 162	0 класс – 626 314 1 класс – 4 198 2 класс – 8 975 3 класс – 26 118 4 класс – 49 784

Таблица 8. Результаты балансировки обучающей выборки «Набора данных 2»

Класс	Число объектов после балансировки	Использованная техника
0 класс	116 162	уменьшено число образцов
1 класс	19 590	увеличено число образцов
2 класс	60 943	увеличено число образцов
3 класс	60 943	без изменений
4 класс	116 162	без изменений

Ансамблевые методы машинного обучения. Ансамблевые методы являются мощным инструментом машинного обучения и находят широкое применение. Цель ансамблевых методов – объединить прогнозы нескольких базовых оценок, построенных с заданным алгоритмом обучения, чтобы улучшить обобщаемость / надежность по сравнению с одной оценкой. Модели нулевого уровня (*базовые модели*) – это модели, обученные на тренировочных данных, прогнозы которых собираются. Модель первого уровня (*метамодель*) – это модель, которая учится наилучшим образом комбинировать прогнозы базовых моделей.

В методах усреднения главный принцип состоит в том, чтобы построить несколько оценщиков независимо, а затем усреднить их прогнозы. Комбинированная оценка обычно лучше, чем любая из оценок с одной базой, потому что ее дисперсия уменьшается. В методах усиления базовые оценки строятся последовательно, и каждый пытается уменьшить смещение комбинированной оценки. Мотивация состоит в том, чтобы объединить несколько слабых моделей для создания мощного ансамбля. Наиболее широко используются следующие ансамблевые методы: бэггинг (модельное усреднение), бустинг (модельное усиление) и стэкинг (модельное наложение) [6-9]. В рамках реализации этапа проекта было рассмотрено два метода построения ансамблей моделей: бэггинг (модельное усреднение) и мягкое и жесткое голосование.

Бэггинг (*bagging, bootstrap aggregation*), или модельное усреднение, – это разновидность параллельного ансамблирования, которая используется для урегулирования проблем, связанных с высокой дисперсией в моделях машинного обучения. Бутстраповская часть бэггинга относится к наборам данных, используемым для тренировки членов ансамбля. В частности, если существует k подмоделей, то для тренировки каждой подмодели ансамбля используется k отдельных наборов данных. Каждый набор данных строится путем случайного отбора экземпляров из изначального тренировочного набора данных (с возвратом). Это означает, что существует высокая вероятность того, что в любом из k наборов данных будут отсутствовать некоторые тренировочные примеры, но и любой набор данных, вероятно, будет иметь повторяющиеся тренировочные примеры. Агрегирование выполняется на результате работы многочисленных членов ансамблевой модели – берется либо среднее значение в случае регрессионной задачи, либо большинство голосов в случае классификационной [10-14].



Рисунок 7. Схематичное представление бэггинга

Жесткое голосование и мягкое голосование – это два разных способа объединения предсказаний нескольких классификаторов в ансамблевых алгоритмах машинного обучения. При жестком голосовании каждый классификатор делает свое предсказание, а предсказание ансамбля – это просто большинство голосов. Мягкое голосование более точное, но и более сложное в реализации. Выбор оптимального подхода зависит от конкретного приложения.

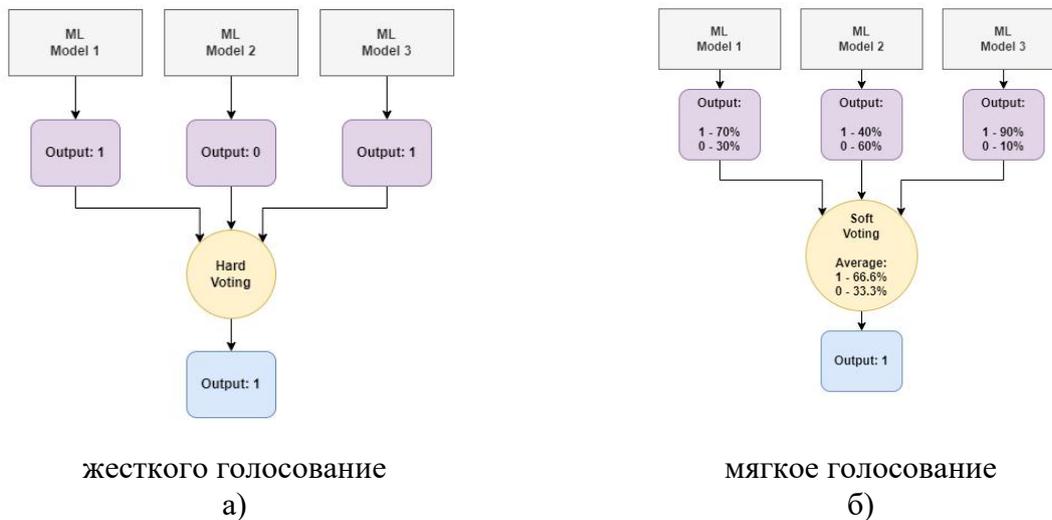


Рисунок 8. Схематичное представление а) – жесткого и б) – мягкого голосования

Рассмотренные методы были использованы при реализации ансамблей нейронных сетей. Искусственные нейронные сети являются современным подходом машинного обучения. Нейронные сети находят эффективное применение в задачах анализа различных типов данных: изображений, видеорядов, текстовых и аудио данных, в том числе в задачах классификации данных временных рядов различной природы [15]. Базовая модель нейронной сети была предложена на предыдущих этапах исследования и использована при построении ансамблей нейронных сетей. Базовая модель нейронной сети является гибридной классификационной моделью, включающей свёрточные и полносвязные слои, слои субдекретизации также блоки *GRU (Gated Recurrent Units)*, использованы техники регуляризации [16-18]. Обозначим базовую модель как «*CNN+GRU+Dense* архитектура».

Разработка ансамблей нейронных сетей для классификации режимов функционирования подсистем КА. В случае решения задачи 2х классовой классификации выход нейронной сети содержит 1 нейрон и используется сигмоидная функция активации.

Здесь и далее в качестве оптимизатора был использован алгоритм *Adam*. Результаты этого оптимизатора, как правило, лучше, чем у любого другого алгоритма оптимизации, занимают меньшее время и требуют меньше параметров для настройки. Благодаря этому метод *Adam* рекомендуется в качестве оптимизатора по умолчанию для большинства приложений обучения современных моделей нейронных сетей. Общее число тренировочных объектов, представленных в одном мини-батче равно 1 024, число эпох обучения 100. В процессе обучения использовались механизмы мониторинга модели с использованием метрики точности, а также механизм ранней остановки с целью исключения переобучения модели.

Для реализации нейронных сетей и механизмов их обучения использовались язык *Python*, фреймворк глубокого обучения *Keras*, а также библиотека машинного обучения *sklearn*. При реализации ансамблей нейронных сетей использовалась библиотека *SciKeras*.

Для оценки эффективности применения ансамблевых методов в начале была реализована и обучена базовая модель, а затем эта модель использовалась при реализации ансамблевых моделей методом бэггинга и методом жёсткого голосования. Во всех экспериментах число базовых моделей в ансамбле равно 3. Результаты тестирования (классификационные отчеты) моделей на тестовом наборе из «Набора данных 1» представлены в таблице 9.

Таблица 9. Результаты тестирования (классификационные отчеты) моделей на тестовом наборе из «Набора данных 1»

Архитектура/ансамбль	Классификационный отчет для тестового набора				
	precision	recall	f1-score	support	
Одна « <i>CNN+GRU+Dense</i> архитектура»	0	0.99	0.82	0.90	633920
	1	0.40	0.96	0.56	81469
	accuracy			0.83	715389
	macro avg	0.70	0.89	0.73	715389
	weighted avg	0.93	0.83	0.86	715389
<i>Voting «hard»</i> три « <i>CNN+GRU+Dense</i> архитектуры»	0	0.99	0.82	0.90	633920
	1	0.41	0.97	0.58	81469
	accuracy			0.84	715389
	macro avg	0.70	0.89	0.74	715389
	weighted avg	0.93	0.84	0.86	715389
Bagging три « <i>CNN+GRU+Dense</i> архитектуры»	0	0.99	0.77	0.87	633920
	1	0.35	0.97	0.52	81469
	accuracy			0.79	715389
	macro avg	0.67	0.87	0.69	715389
	weighted avg	0.92	0.79	0.83	715389

В результате экспериментов установлено, что ансамбль, реализованный на основе метода жёсткого голосования позволяет увеличить точность (accuracy) классификации на 1% в сравнение с базовой моделью, с 83% до 84% соответственно для двухклассовой задачи классификации.

В случае решения задачи 5ти классовой классификации выход нейронной сети содержит 5 нейронов и используется функция активации *softmax*. Для оценки

эффективности применения ансамблевых методов в начале была реализована и обучена базовая модель, а затем эта модель использовалась при реализации ансамблевых моделей методом бэггинга и методом мягкого голосования. Во всех экспериментах число базовых моделей в ансамбле равно 3. Результаты тестирования (классификационные отчеты) моделей на тестовом наборе из «Набора данных 2» представлены в таблице 10.

Таблица 10. Результаты тестирования (классификационные отчеты) моделей на тестовом наборе из «Набора данных 2»

Архитектура/ансамбль	Классификационный отчет для тестового набора				
Одна «CNN+GRU+Dense архитектура»		precision	recall	f1-score	support
	0	0.99	0.79	0.88	626314
	1	0.62	0.70	0.66	4198
	2	0.19	0.58	0.28	8975
	3	0.18	0.60	0.28	26118
	4	0.47	0.93	0.63	49784
	accuracy			0.79	715389
	macro avg	0.49	0.72	0.54	715389
weighted avg	0.91	0.79	0.83	715389	
Voting «soft» три «CNN+GRU+Dense архитектуры»		precision	recall	f1-score	support
	0	0.99	0.79	0.88	626314
	1	0.61	0.72	0.66	4198
	2	0.20	0.68	0.31	8975
	3	0.22	0.58	0.32	26118
	4	0.45	0.97	0.61	49784
	accuracy			0.80	715389
	macro avg	0.49	0.75	0.56	715389
weighted avg	0.91	0.80	0.83	715389	
Bagging три «CNN+GRU+Dense архитектуры»		precision	recall	f1-score	support
	0	0.99	0.78	0.87	626314
	1	0.62	0.69	0.66	4198
	2	0.19	0.62	0.29	8975
	3	0.19	0.58	0.29	26118
	4	0.43	0.96	0.60	49784
	accuracy			0.78	715389
	macro avg	0.48	0.72	0.54	715389
weighted avg	0.91	0.78	0.82	715389	

В результате экспериментов установлено, что ансамбль, реализованный на основе метода мягкого голосования позволяет увеличить точность (*accuracy*) классификации на 1% в сравнение с базовой моделью, с 79% до 80% соответственно для пятиклассовой задачи классификации.

Заключение. В результате исследований выполнен разведочный анализ и подготовка наборов данных для решения задач классификации состояний подсистемы КА. Рассмотрены и реализованы ансамблевые методы машинного обучения (бэггинг и голосование) для решения задач классификации состояния подсистем КА (2 класса и 5 классов). Установлено, что ансамбль нейронных сетей, реализованный на основе метода жёсткого голосования позволяет увеличить точность (*accuracy*) классификации (2 класса) состояния подсистемы КА 1% в сравнение с базовой моделью, с 83% до 84% соответственно. Определено, что ансамбль, реализованный на основе метода мягкого голосования, позволяет увеличить точность (*accuracy*) классификации (5 классов) режимов функционирования подсистемы КА на 1% в сравнение с базовой моделью, с 79% до 80% соответственно. Полученные результаты носят прикладной характер и являются элементом комплексной автоматизации и интеллектуализацию процессов оценивания и многомодельного анализа данных телеметрической информации МКА.

Дальнейшие исследования целесообразно продолжить в части модификации использованной и поиска новых эффективных архитектур нейронных сетей для

построения ансамблей и инжиниринга новых признаков, описывающих состояние подсистемы КА в виду их ограниченного числа.

Список литературы

- [1] Концепция проактивного управления сложными объектами: теоретические и технологические основы / М. Ю. Охтилев [и др.] // Известия вузов. Приборостроение. – 2014. – Т. 57, №11. – С.7–14.
- [2] Проект стратегии информационных технологий Госкорпорации «Роскосмос». – [Электронный документ]. – Режим доступа: <https://www.roscosmos.ru/25892/> (дата обращения: 07.06.2023).
- [3] CSMOUTE: Combined Synthetic Oversampling and Undersampling Technique for Imbalanced Data Classification [Электронный документ]. -- Режим доступа: <https://arxiv.org/pdf/2004.03409.pdf> (дата обращения: 02.11.2023).
- [4] Deep Synthetic Minority Over-Sampling Technique [Электронный документ]. -- Режим доступа: <https://arxiv.org/ftp/arxiv/papers/2003/2003.09788.pdf> (дата обращения: 04.12.2023).
- [5] Automatic Feature Selection [Электронный документ]. – Режим доступа: // <https://amueller.github.io/aml/05-advanced-topics/12-feature-selection.html> (дата обращения: 04.12.2023).
- [6] Брейман Л. Вставка небольших голосов для классификации в больших базах данных и в Интернете / Л. Брейман // Машинное обучение. – 1999. – Том 36, № 1. – С. 85-103.
- [7] Брейман Л. Предикторы упаковки / Л. Брейман // Машинное обучение. – 1996. – Том 24, № 2. – С. 123-140.
- [8] Хо Т. Метод случайных подпространств для построения лесов решений / Т. Хо // Анализ шаблонов и машинный интеллект. – 1998. – Том 20, № 8. – С. 832-844.
- [9] Луппе Г. Ансамбли на случайных участках / Г. Луппе и П. Геуртс // Машинное обучение и обнаружение знаний в базах данных. – 2012. – С. 346-361.
- [10] Tibshirani R. The Elements of Statistical Learning / Tibshirani, R., Friedman, J. – 2001. – Springer, New York. – 764 p.
- [11] Domingos P. A Unified Bias-Variance Decomposition and its Applications / P. Domingos // Proc. 17th International Conf. on Machine Learning. – 2000. – PP. 231-238.
- [12] Breiman L. Random Forests / Breiman L. // Machine Learning. – 2001. – Vol. 451. PP. 5–32.
- [13] Власенко М. Ансамблевые методы машинного обучения в скоринге розничных клиентов [Электронный документ]. -- Режим доступа: <https://www.nbrb.by/bv/pdf/articles/11056.pdf> (дата обращения: 20.10.2023).
- [14] Вольперт Д.Х. Сложное обобщение / Дэвид Х.Вольперт // Нейронные сети. – 1992. – Том 5, №2. – С. 241-259.
- [15] Fawaz H. I. et al. (2019) Muller Deep learning for time series classification: a review. Data Mining and Knowledge Discovery. 33, 917–963.
- [16] Skobtsov V.Y. Automatic Searching the Neural Network Models for Time Series Classification of Small Spacecraft's Telemetry Data with Genetic Algorithms / Skobtsov, V.Y., Stasiuk, A. // Artificial Intelligence Application in Networks and Systems. CSOC 2023. Lecture Notes in Networks and Systems. – 2023. – Vol. 724. – PP. 800–811. doi.org/10.1007/978-3-031-35314-7_66.
- [17] Skobtsov V.Yu. Binary classification of small satellites telemetry data based on deep learning approach // Applied Aspects of Information Technology. 2021. Vol. 4 No.4, pp. 299–310.
- [18] В.Ю.Скобцов, Б.В.Соколов Гибридные нейросетевые модели в задаче мультиклассовой классификации данных телеметрической информации малых космических аппаратов // Вестник ВГУ. Системный анализ и информационные технологии, 2022, № 3, С.99-114

Авторский вклад

Лукашевич Марина Михайловна – разработка нейросетевых классификаторов на основе ансамблевых методов, подготовка и балансировка наборов данных.

Скобцов Вадим Юрьевич – разработка базовой нейросетевой модели, сравнение результатов классификации, анализ полученных результатов.

Инютин Александр Владимирович – руководитель проекта, постановка задачи исследования.

Ганченко Валентин Вячеславович – разведочный анализ данных.

NEURAL NETWORKS ENSEMBLES FOR CLASSIFICATION OF OPERATION STATES AND MODES OF SPACECRAFT SUBSYSTEMS

M.M. Lukashevich
*Associate Professor,
Department of
Information Systems
Management, BSU, PhD
of Technical sciences,
Associate Professor*

V.Yu. Skobtsov
*Associate Professor,
Department of
Computer Technology
and Software
Engineering, Saint
Petersburg State
University of Aerospace
Instrumentation, PhD,
Associate Professor*

A.V. Inyutin
*Head of Laboratory,
UIIP*

V.V. Ganchenko
*Researcher, UIIP, PhD
of Technical sciences*

Abstract. Exploratory analysis and preparation of datasets for solving spacecraft subsystem state and modes classification problems are performed. Ensemble methods of machine learning for solving the tasks of classifying the state and mode of spacecraft subsystems are considered and realized. The obtained results are of applied nature and are an element of complex automation and intellectualization of the processes of evaluation and analysis of small spacecraft telemetry data.

Keywords: neural networks, ensemble methods, classification, on-board equipment, telemetry information, small spacecraft.