

АНСАМБЛЕВЫЕ МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ УГЛОВ ЭЙЛЕРА В ИНЕРЦИАЛЬНОЙ НАВИГАЦИОННОЙ СИСТЕМЕ

А.Н. Греков^{1,2}, А.А. Кабанов²

¹Институт природно-технических систем,
РФ, г. Севастополь, ул. Ленина, 28

²Севастопольский государственный университет,
РФ, г. Севастополь, ул. Университетская, 33
E-mail: i@angrekov.ru

Работа посвящена повышению достоверности навигационной информации автономных платформ, применяемых для изучения океанов и морей, а именно: определению углов Эйлера с использованием экспериментальных данных, формирующихся на выходе инерциальной навигационной системы, построенной на основе MEMS датчиков. Были рассмотрены два ансамблевых метода машинного обучения: мажоритарное голосование (голосование большинством) и взвешенное большинство голосов. Ансамбли были сформированы путём объединения трёх методов обучения с учителем: метода опорных векторов (support vector machine – SVM), k-ближайших соседей (k-nearest neighbors – KNN) и дерева принятия решений. Выполнена оптимизация гиперпараметров трёх этих классификаторов. В результате объединения оптимизированных классификаторов в ансамбль со взвешенным большинством голосов получен прирост правильности (ассигасу) классификации по сравнению с индивидуальными алгоритмами: 0,92 на обучающем и тестовом наборах данных.

Ключевые слова: ансамблевые методы, дерево решений, k-ближайших соседей, машинное обучение, инерциальная навигационная система, автономные подводные аппараты

Поступила в редакцию: 22.12.2021. После доработки: 18.02.2022.

Введение. Многие страны мира проявляют огромный интерес к океану, который обладает богатыми ресурсами и энергией. В связи с этим ведутся активные исследования и разработка большого количества морского оборудования для изучения и освоения океана [1]. Наряду с классическими методами исследования океанов в последние десятилетия активно внедряются автономные платформы, такие как автономные необитаемые подводные аппараты (АНПА) [2], глайдеры [3] и автономные надводные миникорабли [4].

Для выполнения широкомасштабных съемок с приемлемой точностью таким платформам необходима достоверная навигационная информация, получаемая с помощью современных навигационных систем, к числу которых относятся инерциальные навигационные системы [5].

Развитие инерциальных навигацион-

ных систем связано со многими факторами: повышение точности, надежности, стабильности, компактности первичных датчиков и уменьшение их стоимости. Немаловажную роль в развитии играет совершенствование методов и разработка устойчивых алгоритмов обработки навигационной информации.

В настоящее время набирают всё большую популярность современные методы, улучшающие навигационные параметры с применением технологий машинного обучения [6]. В предлагаемой работе мы исследовали применение ансамблевых методов машинного обучения для определения углов Эйлера с использованием инерциальной навигационной системы на основе MEMS датчиков.

Материалы и методы. Целью ансамблевых методов является объединение прогнозов нескольких классификаторов, построенных с заданным алго-

ритмом обучения, для улучшения обобщаемости (надежности) по сравнению с оценкой индивидуального классификатора [7]. В данной работе мы применим два метода объединения: мажоритарное голосование (голосование большинством) [8] и взвешенное большинство голосов [9]. При голосовании большинством прогнозируемая метка класса для конкретной выборки является меткой класса, которая была спрогнозирована большинством классификаторов, предсказываемых каждым отдельным классификатором [10]. Термин большинство голосов относится только к конфигурациям с двоичными классами. Но, принцип мажоритарного голосования легко обобщить на многоклассовые конфигурации, в итоге получив, так называемое, голосование относительным большинством голосов. В таком случае мы выбираем метку класса, которая получила наибольшее число голосов. Такой классификатор полезен для набора одинаково хорошо работающих моделей, чтобы уравновесить их индивидуальные недостатки [11].

В отличие от мажоритарного голосования взвешенное большинство голосов возвращает метку класса как аргумент максимизации суммы предсказанных вероятностей. Каждому классификатору присваивается определенные веса с помощью параметра весов. Когда веса получены, прогнозируемые вероятности классов для каждого классификатора собираются, умножаются на вес классификатора и усредняются. Окончательная метка класса затем получается из метки класса с наивысшей средней вероятностью.

В математической форме можно записать взвешенное большинство голосов следующим образом:

$$\mathcal{C} = \arg \max \sum_{j=1}^m w_j \chi_A(C_j(x) = i), \quad (1)$$

где w_j – вес, ассоциированный с базовым классификатором C_j ; \mathcal{C} – метка класса,

спрогнозированная ансамблем; χ_A – характеристическая или индикаторная функция, которая возвращает 1, если спрогнозированный класс j -го классификатора соответствует i ($C_j(x) = i$). Для равных весов уравнение можно упростить, записав его так:

$$\mathcal{C} = \text{мода}\{C_1(x), C_2(x), \dots, C_m(x)\}. \quad (2)$$

Для реализации этих двух методов объединения классификаторов воспользуемся модулем ensemble библиотеки scikit-learn [12] с параметром voting класса VotingClassifier «soft» для взвешенного большинства голосов и «hard» для мажоритарного голосования.

В качестве входных данных для классификаторов используется информация, получаемая с MEMS датчиков разработанной навигационной платформы. Состав платформы и метод получения данных подробно рассмотрены в работах [13, 14].

Методы взвешенного большинства голосов и мажоритарного голосования требуют минимум два классификатора. В работе мы реализовали три наиболее распространенных метода обучения с учителем: метод опорных векторов (support vector machine – SVM), k -ближайших соседей (k -nearest neighbors – KNN) и дерево принятия решений.

Метод опорных векторов подробно исследован в работе [13]. Там же по кривым проверки были получены оптимальные значения гиперпараметров для SVM классификатора (ядро linear и обратный параметр регуляризации $C = 10$).

Оптимизацию гиперпараметров остальных моделей машинного обучения для подготовленного набора данных проведем методом поиска по сетке [15], реализованным в модуле model_selection библиотеки scikit-learn [12], с перекрестной проверкой по 10 блокам [16, 17]. В качестве метрики эффективности моделей применим долю правильных ответов алгоритма (accracy). В русскоязычной литературе нет устоявшегося перевода

термина «ассигасу» в контексте машинного обучения. Воспользуемся одним из популярных вариантов «Правильность» [18].

KNN. Алгоритм KNN для классификации каждой новой точки осуществляет поиск k ближайших соседних точек в обучающем наборе. Новая точка получает ту метку, которой помечено большинство соседей [19–21]. В качестве меры удаленности новой точки от ближайших соседей в KNN может использоваться евклидово расстояние:

$$d_{\text{евклидово}} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

и манхэттенское расстояние:

$$d_{\text{манхэттенское}} = \sum_{i=1}^n |x_i - y_i|. \quad (4)$$

По умолчанию класс NearestNeighbors использует расстояние Минковского:

$$d_{\text{Минковского}} = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{1/p}, \quad (5)$$

где x_i и y_i – два наблюдения, между которыми мы вычисляем расстояние. Расстояние Минковского включает гиперпараметр p , где $p = 1$ – это манхэттенское расстояние, $p = 2$ – евклидово расстояние и т.д. В библиотеке scikit-learn по умолчанию $p = 2$.

Дерево принятия решений. Дерево принятия решений является непараметрическим методом обучения с учителем, позволяющим разбить набор данных на

ветви с последующим принятием простых решений на каждом уровне [22, 23]. Окончательное решение получают, спускаясь по ветвям этого дерева. Обучение дерева принятия решений – систематический метод построения прогностической модели при классификации, основанный на данных. В работе Лео Бреймана с коллегами [24] заложены теоретические основания деревьев принятия решений, используемые и по сей день.

Результаты. Перед реализацией классификаторов проведено разделение набора данных на обучающий и испытательный наборы, а также стандартизация признаков аналогично [13]. Применим метод поиска по сетке для оптимизации гиперпараметров классификатора KNN. Для этого необходимо задать перечень и диапазоны гиперпараметров (напомним, что в качестве метрики эффективности выступает доля правильных ответов). Как отмечалось выше, KNN может определять расстояние от ближайших соседей через евклидово, манхэттенское расстояние и расстояние Минковского (табл. 1). Кроме того, огромное значение имеет количество точек k , до которых мы определяем расстояние. Как правило, более высокие значения k уменьшают влияние шума на классификацию [25], но делают границы между классами менее четкими. Также для прогноза может использоваться разная весовая функция: однородные веса (все точки в каждой области имеют одинаковый вес) и вес точек, обратный расстоянию до них (ближайшие соседние точки к искомой будут иметь большее влияние, чем точки, находящиеся дальше).

Таблица 1. Перечень и диапазон основных гиперпараметров классификатора KNN

Гиперпараметр	Диапазон
Метрика расстояния	Манхэттенское, евклидово и Минковского ($p = 3$)
Количество соседей	5, 10, 20, 50, 100
Весовая функция	Однородная и зависящая от расстояния

В результате поиска по сетке были получены низкие значения правильности (ассигасу) классификации 0,52 (рис. 1) при оптимальных значениях гиперпараметров, представленных в табл. 2. При дальнейшем увеличении количества соседей правильность (ассигасу) классификации убывает. Низкое значение пра-

вильности (ассигасу) классификации говорит о том, что необходима дополнительная обработка данных перед классификацией (например, понижение размерности), но это требует дополнительных исследований и выходит за рамки данной работы.

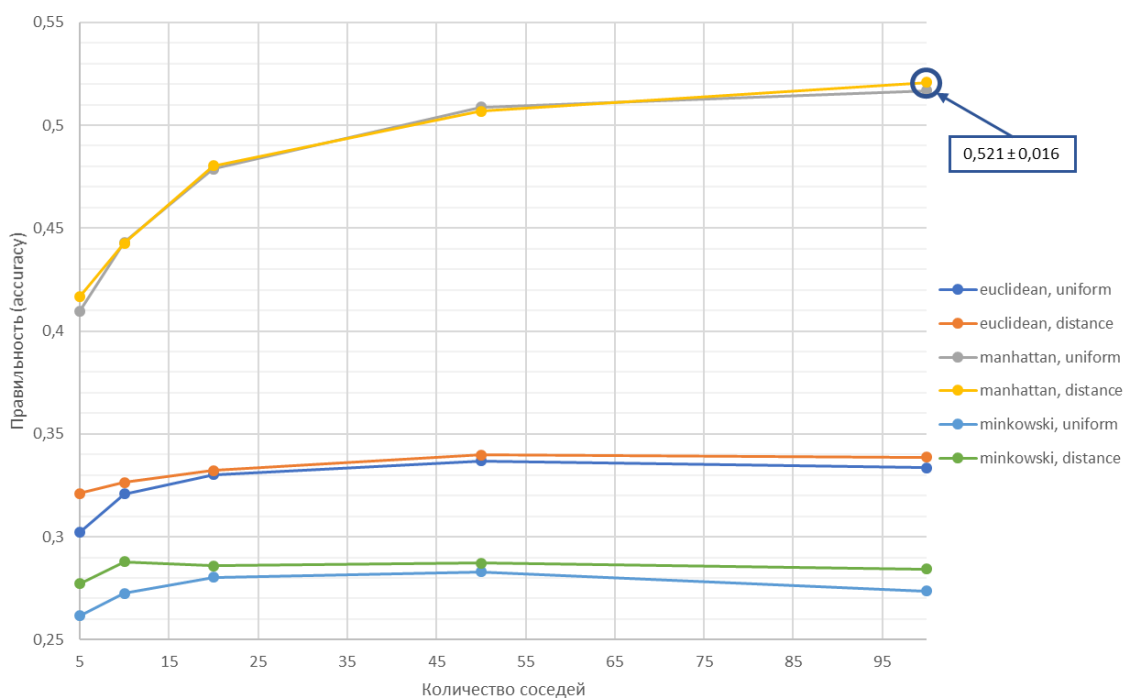


Рис. 1. Зависимость правильности классификации (ассигасу) алгоритма KNN с нормализацией признаков от количества соседей при различных метриках расстояния и весовых функциях

Fig. 1. Classification accuracy dependency of the KNN algorithm with feature normalization on the number of neighbors for different distance metrics and weight functions

Таблица 2. Оптимальные значения гиперпараметров алгоритма KNN с нормализацией признаков

Гиперпараметр	Диапазон
Метрика расстояния	Манхэттенское
Количество соседей	100
Весовая функция	Зависящая от расстояния

В нашем случае хорошие результаты классификации были получены на данных без предварительной стандартизации: правильность (ассигасу) класси-

фикации 0,889 (рис. 2) при значениях гиперпараметров, представленных в табл. 3.

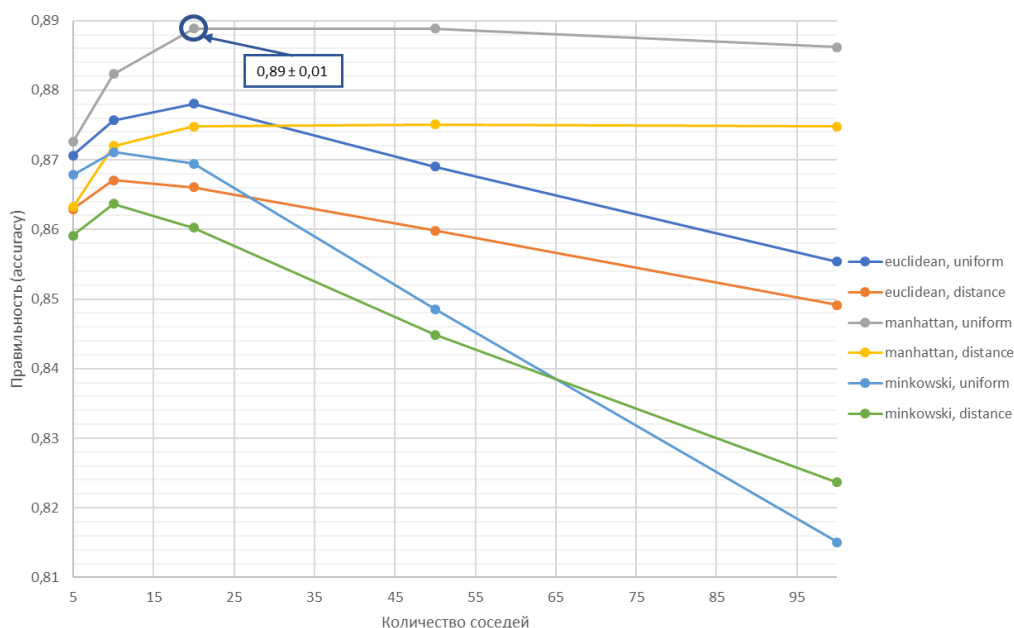


Рис. 2. Зависимость правильности классификации (ассигасу) алгоритма KNN без нормализации признаков от количества соседей при различных метриках расстояния и весовых функциях

Fig. 2. Classification accuracy dependency of the KNN algorithm without feature normalization on the number of neighbors for different distance metrics and weight functions

Таблица 3. Оптимальные значения гиперпараметров алгоритма KNN без нормализации признаков

Гиперпараметр	Диапазон
Метрика расстояния	Манхэттенское
Количество соседей	20
Весовая функция	Однородная

Для реализации поиска по сетке для дерева решений воспользуемся перечнем и диапазоном гиперпараметров (табл. 4).

Таблица 4. Перечень и диапазон гиперпараметров дерева решений

Гиперпараметр	Диапазон
Критерий разбиения	Джини и энтропия
Максимальная глубина дерева	1 – 15
Минимальное количество выборок для конечного узла (msl)	1 – 4

Правильность (ассигасу) классификации, полученная при оптимальных значениях гиперпараметров (табл. 5), составляет $0,91 \pm 0,01$ (рис. 3). Получен-

ные результаты показывают, что эффективность прогнозирования индивидуальных классификаторов почти равна.

Таблица 5. Оптимальные значения гиперпараметров для дерева решений

Гиперпараметр	Диапазон
Критерий разбиения	Энтропия
Максимальная глубина дерева	6
Минимальное количество выборок для конечного узла (msl)	1

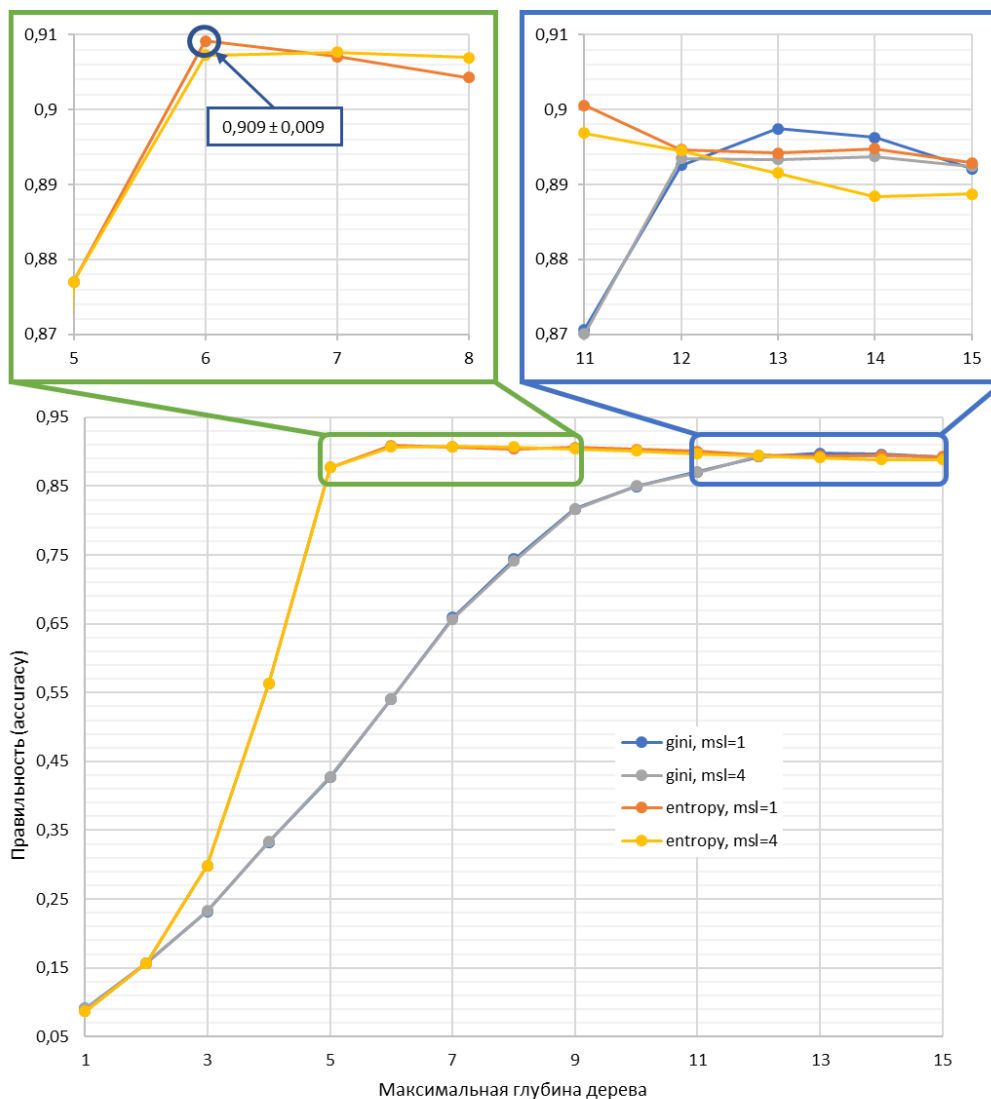


Рис. 3. Зависимость правильности классификации (ассигасу) дерева решений от максимальной глубины дерева при различных значениях минимального количества выборков для конечного узла (msl)

Fig. 3. Classification accuracy dependency of the decision tree on the maximum depth of the tree for various minimum number of samples in a leaf (msl)

Теперь объединим три наших классификатора с оптимальными гиперпараметрами. В качестве метрики эффективности нашего ансамбля также воспользуемся правильностью (ассигасу) классификации, а валидацию проведем с помощью перекрестной проверки по 10 блокам. Для взвешенного большинства голосов правильность (ассигасу) классификации обучающего набора составила $0,92 \pm 0,01$. При этом правильность (ассигасу) классификации на тестовом наборе показала следующие результаты:

- SVM 0,91,

- KNN 0,90,
- Дерево решений 0,91,
- Ансамбль 0,92.

Для мажоритарного голосования правильность (ассигасу) классификации обучающего набора составила $0,91 \pm 0,01$, а правильность (ассигасу) классификации на тестовом наборе показала следующие результаты:

- SVM 0,91,
- KNN 0,90,
- Дерево решений 0,91,
- Ансамбль 0,91.

Выводы. Используя данные, полученные с MEMS датчиков разработанной навигационной платформы, были опробованы две модели классификатора: KNN и дерево решений. После оптимизации гиперпараметров правильность (ассигасу) классификации модели KNN составила $0,89 \pm 0,01$, а модели дерево решений – $0,91 \pm 0,01$. Проверка на тестовом наборе данных также подтвердила хорошую эффективность этих классификаторов: правильность (ассигасу) KNN – $0,90$, а дерева решений – $0,91$. Объединение двух вышеупомянутых классификаторов и классификатора SVM в ансамбль со взвешенным большинством голосов дает небольшой прирост правильности (ассигасу) классификации: $0,92$ на обучающем и тестовом наборах данных, что позволяет рекомендовать полученные результаты исследований для улучшения качества навигационной информации. В случае ансамбля с мажоритарным голосованием значимого прироста правильности (ассигасу) классификации не выявлено по сравнению с индивидуальными классификаторами.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации, Соглашение Минобрнауки РФ № 075-03-2021-092/5 от 29.09.2021, FEFM-2021-0014 № 121111600136-3.

СПИСОК ЛИТЕРАТУРЫ

1. Греков А.Н., Греков Н.А., Сычев Е.Н. Среднечастотные акустические методы и средства для исследования водной среды. Севастополь: ИПТС, 2020. 126 с. ISBN 978-5-6044196-6-3
2. Киселев Л.В., Инзарцев А.В., Матвиенко Ю.В. Создание интеллектуальных АНПА и проблемы интеграции научных исследований // Подводные исследования и робототехника. 2006. № 1. С. 6–17.
3. Гайкович Б.А., Занин В.Ю., Кожемякин И.В. Вопросы разработки морских робототехнических платформ на примере создания подводного аппарата типа «Глайдер» // Морская робототехника. Перспективные системы и задачи управления: труды конференции. 2016. С. 151.
4. Шишкин Ю.Е., Греков А.Н. Концепция интеллектуальной системы автоматизированного экологического мониторинга на базе малогабаритных автономных роботов // Системы контроля окружающей среды. 2018. № 4(34). С. 63–69.
5. Греков А.Н., Греков Н.А., Алексеев С.Ю. Бесплатформенный навигационный комплекс с инерциальной системой ориентации на «грубых» чувствительных элементах и способ коррекции его инерциальных датчиков; Пат. 2548115 Россия, МПК G01C 23/00. № 2014151906/93; заявл. 18.12.14; опубл. 10.04.15; Бюл. № 10.
6. Li Y. et al. Inertial Sensing Meets Machine Learning: Opportunity or Challenge? // IEEE Transactions on Intelligent Transportation Systems. 2021.
7. Raschka S., Mirjalili V. Python Machine Learning: Machine Learning and Deep Learning with Python. Scikit-Learn, and TensorFlow. Second edition ed. 2017.
8. Ruta D., Gabrys B. Classifier selection for majority voting // Information fusion. 2005. Vol. 6. № 1. P. 63–81.
9. Kuncheva L.I. Combining pattern classifiers: methods and algorithms. John Wiley & Sons, 2014.
10. Alpaydin E. Introduction to machine learning. MIT press, 2020.
11. Dietterich T.G. Ensemble Methods in Machine Learning. In International Workshop on Multiple Classifier Systems; Springer: Berlin/Heidelberg, Germany, 2000. P. 1–15.
12. Pedregosa F. et al. Scikit-learn: Machine Learning in Python // Journal of Machine Learning Research. 2011. Vol. 12. P. 2825–2830.
13. Греков А.Н., Кабанов А.А., Алексеев С.Ю. Метод опорных векторов для определения углов Эйлера в инерциальной навигационной системе // Системы контроля окружающей среды. 2021. № 4 (46). С. 134–142. DOI:10.33075/2220-5861-2021-4-134-142
14. Греков А.Н., Алексеев С.Ю., Башикиров В.Ю. Результаты лабораторных испытаний подводной навигационной

системы для аппаратов экологического контроля // Системы контроля окружающей среды. 2020. № 3 (41). С. 65–74. DOI:10.33075/2220-5861-2020-3-65-74

15. *Molin S., Jee K.* Hands-On Data Analysis with Pandas - Second Edition: A Python Data Science Handbook for Data Collection, Wrangling, Analysis, and Visualization. Packt Publishing, 2021. 788 p. ISBN 9781800563452.

16. *Kohavi R. et al.* A study of cross-validation and bootstrap for accuracy estimation and model selection // Ijcai. 1995. Vol. 14. № 2. P. 1137–1145.

17. *Varma S., Simon R.* Bias in error estimation when using cross-validation for model selection // BMC bioinformatics. 2006. Vol. 7. № 1. P. 1–8.

18. *Раиска С., Мурджалили В.* Python и машинное обучение: машинное и глубокое обучение с использованием Python, scikit-learn и TensorFlow 2. СПб.: ООО «Диалектика», 2020. 848 с.

19. *Peterson L.E.* K-nearest neighbor // Scholarpedia. 2009. Vol. 4. № 2. P. 1883.

20. *Zhang S. et al.* Efficient kNN classification with different numbers of nearest neighbors // IEEE transactions on neural networks and learning systems. 2017. Vol. 29. № 5. P. 1774–1785.

21. *Guo G. et al.* KNN model-based approach in classification // OTM Confederal International Conferences "On the Move to Meaningful Internet Systems". Springer, Berlin, Heidelberg, 2003. P. 986–996.

22. *Джоши П.* Искусственный интеллект с примерами на Python. М.; СПб.: Диалектика. 2019.

23. *Quinlan J.R.* Induction of decision trees // Machine learning. 1986. Vol. 1. № 1. P. 81–106.

24. *Breiman L., Friedman J.H., Olshen R.A., Stone C.J.* Classification and regression trees. Wadsworth Inc. 1984. Vol. 67. 368 p.

25. *Everitt B.S., Landau S., Leese M., Stahl D.* Miscellaneous Clustering Methods // Cluster Analysis: 5th Edition, John Wiley & Sons, Ltd., Chichester, UK. 2011. 352 p.

ENSEMBLE MACHINE LEARNING METHODS FOR EULER ANGLES DETECTION IN AN INERTIAL NAVIGATION SYSTEM

A.N. Grekov^{1,2}, A.A. Kabanov²

¹Institute of Natural and Technical Systems, RF, Sevastopol, Lenin St., 28

²Sevastopol State University, RF, Sevastopol, Universitetskaya St., 33

E-mail: *i@angrekov.ru*

The work is focused on increasing the reliability of navigation information of autonomous platforms used in studying oceans and seas, namely: determining the Euler angles using experimental data generated at the output of an inertial navigation system built on the basis of MEMS sensors. Two ensemble methods of machine learning are considered: majority voting (voting by the majority) and weighted majority voting. The ensembles are formed by combining three supervised learning methods: support vector machine (SVM), k-nearest neighbors (KNN), and decision trees. Optimization of hyperparameters of these three classifiers is performed. As a result of combining the optimized classifiers into an ensemble with a weighted majority, an increase in classification accuracy (accuracy) is obtained compared to individual algorithms: 0.92 on the training and test data sets.

Keywords: ensemble methods, decision tree, k-nearest neighbors, machine learning, inertial navigation system, autonomous underwater vehicles

REFERENCES

1. *Grekov A.N., Grekov N.A., and Sychov E.N.* Srednechastotnye akusticheskie metody i sredstva dlya issledovaniya vodnoj sredy (Mid-frequency acoustic methods and instruments for the study of the aquatic environment). Sevastopol: IPTS, 2020, 126 p. ISBN 978-5-6044196-6-3

2. *Kiselev L.V., Inzarcev A.V., and Matvienko Yu.V.* Sozдание intellektual'nyh ANPA i problemy integracii nauchnyh issledovanij (Creation of intelligent AUVs and problems of scientific research integra-

tion). *Podvodnye issledovaniya i robototekhnika*, 2006, No. 1. pp. 6–17.

3. *Gajkovich B.A., Zanin V.YU., and Kozhemyakin I.V.* Voprosy razrabotki morskikh robototekhnicheskikh platform na primere sozdaniya podvodnogo apparata tipa "Glajder" (Issues of development of marine robotic platforms on the example of the creation of an underwater vehicle of the "Glider" type). *Morskaya robototekhnika. Perspektivnye sistemy i zadachi upravleniya: trudy konferencii*, 2016, pp. 151.

4. *Shishkin Y.E. and Grekov A.N.* Koncepcija intellektual'noj sistemy avtomatizirovannogo jekologicheskogo monitoringa na baze malogabaritnyh avtonomnyh robotov (The concept of automated environmental monitoring intellectual system based on compact autonomous robots). *Sistemy kontrolja okruzhajushhej sredy*, 2018, No. 4(34), pp. 63–69.

5. *Grekov A.N., Grekov N.A., and Alekseev S.Yu.* Besplatformennyj navigacionnyj kompleks s inercial'noj sistemoy orientacii na "grubyyh" chuvstvitel'nyh elementah i sposob korrekcii ego inercial'nyh datchikov (Strapdown navigation system with an inertial orientation system on "coarse" sensitive elements and a method for correcting its inertial sensors); Pat. 2548115 Rossiya, MPK G01S 23/00. № 2014151906/93; zayavl. 18.12.14; opubl. 10.04.15, Byul. № 10.

6. *Li Y. et al.* Inertial Sensing Meets Machine Learning: Opportunity or Challenge? *IEEE Transactions on Intelligent Transportation Systems*, 2021.

7. *Raschka S. and Mirjalili V.* Python Machine Learning: Machine Learning and Deep Learning with Python. Scikit-Learn, and TensorFlow, 2017.

8. *Ruta D. and Gabrys B.* Classifier selection for majority voting. *Information fusion*, 2005, Vol. 6, No. 1, pp. 63–81.

9. *Kuncheva L.I.* Combining pattern classifiers: methods and algorithms. John Wiley & Sons, 2014.

10. *Alpaydin E.* Introduction to machine learning. MIT press, 2020.

11. *Dietterich T.G.* Ensemble Methods in Machine Learning. In International Workshop on Multiple Classifier Systems; Springer: Berlin/Heidelberg, Germany, 2000, pp. 1–15.

12. *Pedregosa F. et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 2011, Vol. 12, pp. 2825–2830.

13. *Grekov A.N., Kabanov A.A., and Alekseev S.Yu.* Metod opornyh vektorov dlya opredeleniya uglov Ejlera v inercial'noj navigacionnoj sisteme (Support vector machine for determining Euler angles in an inertial navigation system). *Sistemy kontrolja okruzhajushhej sredy*, 2021, No. 4 (46), pp. 134–142. DOI:10.33075/2220-5861-2021-4-134-142

14. *Grekov A.N., Alekseev S.Yu., and Bashkirov V.Yu.* Rezul'taty laboratornyh ispytaniy podvodnoj navigacionnoj sistemy dlya apparatov jekologicheskogo kontrolya (The results of laboratory tests underwater navigation system for environmental monitoring devices). *Sistemy kontrolya okruzhajushhej sredy*, 2020, No. 3 (41), pp. 65–74. DOI:10.33075/2220-5861-2020-3-65-74

15. *Molin S. and Jee K.* Hands-On Data Analysis with Pandas - Second Edition: A Python Data Science Handbook for Data Collection, Wrangling, Analysis, and Visualization. Packt Publishing, 2021, 788 p. ISBN 9781800563452.

16. *Kohavi R. et al.* A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai*, 1995, Vol. 14, No. 2, pp. 1137–1145.

17. *Varma S. and Simon R.* Bias in error estimation when using cross-validation for model selection. *BMC bioinformatics*, 2006, Vol. 7, No. 1, pp. 1–8.

18. *Rashka S. and Mirdzhalili V.* Python i mashinnoe obuchenie: mashinnoe i glubokoe obuchenie s ispol'zovaniem Python, scikit-learn i TensorFlow 2 (Python and machine learning: machine and deep learning using Python, scikit-learn and TensorFlow 2), Saint-Petersburg: OOO "Dialektika", 2020, 848 p.

19. *Peterson L.E.* K-nearest neighbor. *Scholarpedia*, 2009, Vol. 4, No. 2, p. 1883.

20. *Zhang S. et al.* Efficient kNN classification with different numbers of nearest neighbors. *IEEE transactions on neural networks and learning systems*, 2017, Vol. 29, No. 5, pp. 1774–1785.

21. *Guo G. et al.* KNN model-based approach in classification. *OTM Confederated International Conferences "On the Move to Meaningful Internet Systems"*, Springer, Berlin, Heidelberg, 2003, pp. 986–996.

22. *Dzhoshi P.* Iskusstvennyj intellekt s primerami na Python (Artificial Intelligence with Python Examples). Moscow; Saint-Petersburg: Dialektika, 2019.

23. *Quinlan J.R.* Induction of decision trees. *Machine learning*, 1986, Vol. 1, No. 1, pp. 81–106.

24. *Breiman L., Friedman J.H., Olshen R.A., and Stone C.J.* Classification and regression trees. Wadsworth Inc. 1984, Vol. 67, 368 p.

25. *Everitt B.S., Landau S., Leese M., and Stahl D.* Miscellaneous Clustering Methods. *Cluster Analysis*: John Wiley & Sons, Ltd., Chichester, UK, 2011, 352 p.